AD-A075 208 ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT--ETC F/G 17/7
PRINCIPLES AND OPERATIONAL ASPECTS OF PRECISION POSITION DETERM--ETC(U)
JUL 79 C T LEONDE
UNCLASSIFIED AGARD-AG-245 NL

1 OF 5
AD
A075208

LEVEL II

AGARD-AG-245

AGARD-AG-245

AD A075208

AGARD

ADVISORY GROUP FOR AEROSPACE RESEARCH & DEVELOPMENT

7 RUE ANCELLE 92200 NEUILLY SUR SEINE FRANCE

AGARDograph No. 245

Principles and Operational Aspects of Precision Position Determination Systems

NORTH ATLANTIC TREATY ORGANIZATION

DISTRIBUTION AND AVAILABILITY
ON BACK COVER

79-10 17 031

(14) AGARD-AG-245

NORTH ATLANTIC TREATY ORGANIZATION

ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT

(ORGANISATION DU TRAITE DE L'ATLANTIQUE NORD)

AGARDograph No.245

(6) **PRINCIPLES AND OPERATIONAL ASPECTS OF PRECISION POSITION DETERMINATION SYSTEMS,**

Edited by

Dr C.T.Leondes
Engineering Systems Department
School of Engineering and Applied Science
Los Angeles, California 90024

(11) Jul 79 (12) 475

(10) Cornelius T./Leondes

This AGARDograph was prepared at the request of the Guidance and Control Panel of AGARD.

400 043 LB

THE MISSION OF AGARD

The mission of AGARD is to bring together the leading personalities of the NATO nations in the fields of science and technology relating to aerospace for the following purposes:

- Exchanging of scientific and technical information;
- Continuously stimulating advances in the aerospace sciences relevant to strengthening the common defence posture;
- Improving the co-operation among member nations in aerospace research and development;
- Providing scientific and technical advice and assistance to the North Atlantic Military Committee in the field of aerospace research and development;
- Rendering scientific and technical assistance, as requested, to other NATO bodies and to member nations in connection with research and development problems in the aerospace field;
- Providing assistance to member nations for the purpose of increasing their scientific and technical potential;
- Recommending effective ways for the member nations to use their research and development capabilities for the common benefit of the NATO community.

The highest authority within AGARD is the National Delegates Board consisting of officially appointed senior representatives from each member nation. The mission of AGARD is carried out through the Panels which are composed of experts appointed by the National Delegates, the Consultant and Exchange Programme and the Aerospace Applications Studies Programme. The results of AGARD work are reported to the member nations and the NATO Authorities through the AGARD series of publications of which this is one.

Participation in AGARD activities is by invitation only and is normally limited to citizens of the NATO nations.

The content of this publication has been reproduced directly from material supplied by AGARD or the authors.

Published July 1979

ISBN 92-835-1330-4

Printed by Technical Editing and Reproduction Ltd
Harford House, 7–9 Charlotte St, London, W1P 1HD

PREFACE

In the early years of aircraft flight navigational means were primitive and could contribute to hazardous situations in adverse weather conditions. With rapid growth in the military and civil use of aircraft it was inevitable that satisfactory navigational improvements would evolve. Thus, radio aids to navigation such as TACAN, OMEGA, LORAN, VORTAC, DECCA, and others were developed. Additionally, inertial systems displayed a remarkable growth over the past two decades, and now proliferate in military and civil applications. The necessity to develop precision position determination systems led, inevitably, to the TRANSIT satellite system used in both military and civil applications at sea, and relied on the relation between the TRANSIT satellite ephemeris and the time history of the doppler signal received by the surface receiver to determine position with precision.

The next step, though, was inevitable, the development of modern precision position determination systems. These are embodied in several principal major new systems, namely NAVSTAR GPS (Global Positioning Satellite), JTIDS (Joint Tactical Information Distribution System), and several other systems receiving interest and consideration. Basically, these systems are based on highly precise DME/trilateration notions. Of course, the knowledge of the precise distance from three known points by radio DME methods (Phase locked loops, for instance) enables the precise determination of one's own position. As explained in greater detail in this book, such methods rely on a precision clock or time reference with the precision of the time reference dictating the precision of position determination. At the same time, an "ultra" accurate time reference suggests a time reference which may be prohibitively expensive for many operational users, military or civil. Thus, this inaccuracy in a user's clock or time accuracy can be viewed as an additional unknown, and can be solved for or corrected for with a fourth measurement, if this is necessary. Thus, four precise measurements from four precisely known points, that is, a precise knowledge of satellite ephemerides, for example, enable the solution of four unknowns with great precision, 3 distances or ranges and a time correction, and thus, the determination of one's own position with great accuracy. Both GPS and JTIDS, and other such systems rely on these fundamental notions of trilateration, and coupled with today's rapidly and constantly advancing electronics semiconductor technology, in addition to other technologies, make such systems a cost effective reality.

It is worth noting here, in passing, that in NAVSTAR GPS the user can determine his position with great accuracy, but then has to report this through separate means to a larger community, if this is necessary. Wheras, JTIDS has such self-reporting capability as a built-in feature, of course, with implications on the system and the equipment, accordingly. Additionally, with respect to JTIDS there are at least two fundamental schemes for operation under consideration, TDMA (Time Division Multiple Access) and DTDMA (Distributed TDMA). Time alone will necessarily resolve which is the most cost effective scheme in the military operational environment where A J margin is primarily a factor for concern. Again, these issues are treated in greater detail in the book. Also treated in greater detail in the book are the different classes of user or receiver equipment.

Test results on systems such as GPS and JTIDS to date suggest phenomenal accuracy, almost near zero error. The implications then expand well beyond more conventional operational notions in military and civil applications. Thus, the near zero initial errors achievable through these systems have profound implications for PGM's (Precision Guided Munitions) of a wide variety. There are many other profound implications apparent, and many others that will become apparent with the passage of time and the learning process that inevitably follows the operational utilization of such systems.

It is a pleasure to acknowledge the many individuals who made this NATO AGARDograph possible. First of all the aid and support of the previous and present Guidance and Control Panel Chairmen, Mr. Morris A.Ostgaard and Mr. Peter Kant was invaluable. Past and present members of the Guidance and Control Panel provided support, enthusiasm, and many valuable suggestions. An especial note of appreciation is due to the Editor's longtime friend and colleague, Mr. Ronald Vaughn, who carried much of the burden in planning the material for JTIDS. The support of the National Delegate Board for this rather substantial project is also greatly appreciated. Additonally, and very importantly, the outstanding support and very competent administration in this highly complex undertaking provided by the past and current Guidance and Control Panel Executives, Colonel Michael Cavenel and Colonel J.C. de Chassey is gratefully acknowledged. Finally, the contributions of the many co-authors and their willing and enthusiastic participation in this venture is gratefully acknowledged. May this volume provide them ample gratification for their efforts for many years to come.

Cornelius T.Leondes
Editor

CONTENTS

Reference

RADIO NAVIGATION SYSTEMS — CURRENT STATUS

by

James S. Tyler
Fred G. Karkalik
Systems Control, Inc.
1801 Page Mill Road
Palo Alto, California 94304

SUMMARY

The current status of radio navigation includes a number of different operational and developmental systems which offer unprecedented capability to a wide variety of air and marine, civil and military users at varying costs and levels of sophistication; however, this status is likely to change significantly over the next several decades. The advent of the Global Positioning System offers the potential to replace the function of several existing systems. Significant changes are also occurring with regard to the number of users, the type of user and the unique requirements of some users. The current government concerns over the proliferation of radio navigation systems and the large costs of developing and operating redundant navigation systems may result in a smaller number of systems which must satisfy a broader range of user requirements. The promising future of radio navigation can only be realized by a careful planning process which involves all segments of the navigation community and results in decisions on which systems will be supported over what periods of time and for what applications. This planning process includes a quantitative assessment of user requirements in a common format, a comparison of candidate system capabilities to meet user needs, and cost tradeoffs versus user benefits, implementation schedules and performance measures.

This paper addresses the life-cycle costs, user requirements, and current system capabilities. The life-cycle costs of several system scenarios are first presented for a limited segment of civil air users in order to illustrate the large magnitude of the costs involved. Current user requirements are summarized for air and marine civil users. Finally, the technical characteristics and operational status of several current systems are presented, namely, Loran-A, Loran-C, Omega, VLF, Decca, VOR/DME, and TACAN.

1. INTRODUCTION

The art and science of radio navigation has evolved from simple non directional beacons to complex world-wide systems over the last half century. It is now entering an era of unparalleled application to air, sea and land users both military and civil, as well as unprecedented changes due to new systems and technology improvements.

The spectrum of present system capabilities ranges from very low frequency systems such as Omega which provides world-wide coverage at relatively low cost, to Loran and Decca which offer higher accuracy and reliable signals in localized areas, to solid-state VORTAC for improved enroute and terminal air navigation, to limited-coverage transit satellite navigation for marine navigation. The advent of the Global Positioning System (GPS), with its high degree of accuracy and world-wide coverage, provides the potential of replacing the capability now provided by several different systems.

At the same time these major changes are occurring due to the introduction of new navigation systems, significant changes can also be foreseen with regard to the type of users and user requirements. The rapid increase of general aviation in the U.S. and the growing number of helicopter operations is changing the mix of air navigation users. The requirement of these users as well as the emphasis of air carrier operations on energy management will motivate the implementation of two- and three-dimensional area navigation. The increased emphasis on safety and environmental impact in coastal confluences and harbor areas establishes a requirement for more precise and reliable radio navaids to replace the more conventional audio-visual aids and reliance on dead-reckoning in marine navigation. Perhaps the biggest change in the user community will be in the land user segment. The civil applications for vehicle position location may eventually result in one of the largest users of radio navigation systems. The military requirements for world-wide, all-weather, very precise navigation for air, sea, and land navigation are clearly the most demanding of all requirements which have lead to the development of the Global Positioning System.

With the unprecedented rate of changes in both system capabilities and user requirements, there is a potential for a proliferation of navigation systems. In order to control costs and assure compatibility, navigation management planning must be applied on a coordinated international level.

Historically, many of the existing radio navigation systems have been developed for a military requirement and then found wide-spread application in the civil sector. Loran-C, Omega, and now satellite navigation are examples; however, once a radio signal suitable for navigation is broadcast, manufacturers develop and promote user equipment, a user base is established, and it becomes difficult to phase a radio navigation system out. The total life-cycle cost of any radio navigation system, including development, operational, maintenance, and user equipment costs, is very large. When all the current and planned future systems' costs are aggregated, the cost is staggering; hence, a current focus by the U.S. Government is to limit the number of navigation systems to avoid redundant capability. The resulting uncertainty in which systems will be implemented or phased out, when and where, is a major issue which impacts all segments of the navigation community.

The government(s) have the responsibility to operate systems which will satisfy both military and civil users in a cost-effective manner over a particular period of time. This responsibility to improve service to an ever wider class of users with up-to-date technology is bound to result in some changes. Too frequent changes cause excessive user equipment replacement costs. On the other hand, commitment to

a given (fixed) system over a long period of time will result in systems which do not incorporate up-to-date technology and therefore may not provide the most effective service for any given period of time. Typically a long-term system which satisfies many requirements can be foreseen (e.g., satellite navigation) but the long delays of implementing such a system require some interim, albeit less effective, system which will still result in improvements over existing systems. Other considerations in the govenment bodies include how to spread development and operational costs among civil users, if at all; and whether the government should be concerned about total costs (system development, operation, and user equipment costs), not just those costs incurred by the government. Thus the government planning process is one of trying to improve the performance and coverage of navigation systems based on up-to-date technology in a manner which is cost-effective over a period of time (typically 20-30 years). Whereas major emphasis in the past has been primarily to satisfy mission effectiveness requirements of the military, it now appears that the greatly increasing civil user base will play a more significant role in influencing which navigation systems will be supported.

From an equipment manufacturer's point of view, the government planning process results in uncertainties with regard to schedules of system implementation, decisions on which competing systems will exist, and what regulations and standards will be effected. There is little incentive for manufacturers to invest in the development of equipment which has a limited market. Also the early development and sales of user equipment prior to a complete system test and evaluation and standards specification may lead to premature and sometimes unsafe use of radio navigation signals.

From the user point of view the uncertainties caused by the government planning and development process, as well as user equipment availability, cause the greatest penalties of all. Investment decisions on new equipment for ships, aircraft, or land vehicles become difficult due to uncertainties of which systems will be available, when, where, at what cost, and with what functional capabilities; as well as what systems may be required by government regulations. Equipping with current systems which are protected for several years in the future (e.g. VOR/DME for aircraft) is perhaps the safest decision; however, the price that is paid with this decision may be the ultimate existence of non-up-to-date technology and perhaps a degradation of performance or economic penalty when compared to the use of the best technology and latest systems. On the other hand, equipping with new systems prematurely can often result in poor performance due to incomplete system coverage (e.g. Omega), or unproven user equipment; and may also result in equipment which at some later date is obsolete. This situation is further complicated for many users who may require, or be required to have, several different enroute and terminal aids. For example, it is not inconceivable that a tanker operating between the United States and Europe might need Loran-C for U.S. waters, Decca for Rotterdam Harbor, and satellite navigation for high-seas navigation.

The interdependency of all segments of the navigation community clearly requires careful planning if the existing potential of radio navigation is to be realized over the next several decades. The essential steps in such a planning process include: (1) quantifying the present and projected future user requirements for radio navigation, (2) quantifying the capabilities of current and potential future systems and comparing them against the requirements, and (3) agreeing on a time-phased and cost-effective mix of systems which will satisfy all user requirements.

This paper focuses on those elements of radio navigation systems planning which have been identified in the U.S. First a summary of those system costs is presented for a limited segment of the navigation users in order to illustrate the magnitude of the total costs involved. Second, those user requirements are summarized which have been identified to date. Finally, the operational status and usage of those current radio navigation systems are summarized which are not covered in subsequent chapters; namely, Loran-C, Omega, Decca, VOR/DME, TACAN, and non-directional beacons.

2. LIFE-CYCLE COSTS

The future of radio navigation is highly dependent on which systems the government(s) support and/or require. Although the capability of systems to satisfy user requirements is the primary factor in these decisions, system costs play an important role in influencing which systems will be supported over a particular time period in order to avoid redundant system capability. Total system costs including user and government costs for civil and military users are difficult to assess; however, some initial steps have been taken by the U.S. Government to quantify partial life-cycle costs for a limited segment of users. These figures are summarized in this section in order to illustrate the substantial costs which are associated with the development, operation, and use of radio navigation systems.

The primary motivation for the analysis of life cycle costs of radio navigation systems at this time is the potential of a satellite system such as GPS to provide the necessary navigation capability to a broad spectrum of military and civil users. Most of the planning and system development which has resulted in today's radio navigation system has not been influenced by the ultimate existence of satellite navigation; hence, the major cost tradeoffs now involve assessment of total system costs of existing systems (including planned improvements to those systems) vs. the total system costs assuming the existence of satellite navigation.

A recent study for the U.S. Department of Transportation, Federal Aviation Administration [1] has addressed the life-cycle costs of civil air navigation system alternatives when applied to operations in the continental U.S. (CONUS), Alaska, off-shore and oceanic regions. These alternatives include: (1) the continued use of VOR/DME including VORTAC improvements, (2) a transition to differential Omega for Alaska, (3) a transition to Loran-C for all regions except oceanic, (4) a transition to GPS for all regions and the retention of VOR for low cost users in the CONUS and Alaska regions, and (5) a transition to GPS for all civil aviation in all regions.

The cumulative costs through the year 2005 for the five different scenarios are shown in Figure 1. Both user and FAA costs are shown. User life cycle cost impacts are limited to those costs associated with replacing enroute navigation avionics, through both voluntary and mandated retrofitting. The FAA costs shown in Figure 1 were limited to only those incremented expenses required to implement, supplement, operate and/or support a system so as to satisfy those requirements established for U.S. civil

KEY	NAVIGATION SYSTEM ALTERNATIVE
1	BASELINE - RETAIN 2ND GEN VOR/DME, OMEGA USED TO PROVIDE OCEANIC AND OFF-SHORE COVERAGE
2	DIFF. OMEGA IN AKA AND AKA OFF-SHORE, 2ND GEN VOR/DME IN CONUS, OMEGA USED TO PROVIDE OCEANIC AND CONUS OFF-SHORE COVERAGE
3	LORAN-C IN ALL REGIONS EXCEPT OMEGA USED TO PROVIDE OCEANIC COVERAGE
4	GPS IN ALL REGIONS
5	GPS IN ALL REGIONS EXCEPT LOW COST VOR RETAINED IN CONUS & AKA
(shaded)	RANGE OF COSTS ATTRIBUTABLE TO VARIATIONS IN IMPLEMENTATION/TRANSITION SCENARIO SCHEDULES. NOTE: THE MIN (MAX) 'USER + FAA' COSTS DO NOT GENERALLY CORRESPOND TO THE SUM OF THE MIN (MAX) 'SYSTEM USER' AND 'FAA' COSTS, SINCE SCENARIOS THAT PRODUCED MIN (MAX) 'SYSTEM USER' COSTS TYPICALLY DID NOT ALSO PRODUCE MIN (MAX) 'FAA' COSTS AND VICE VERSA.

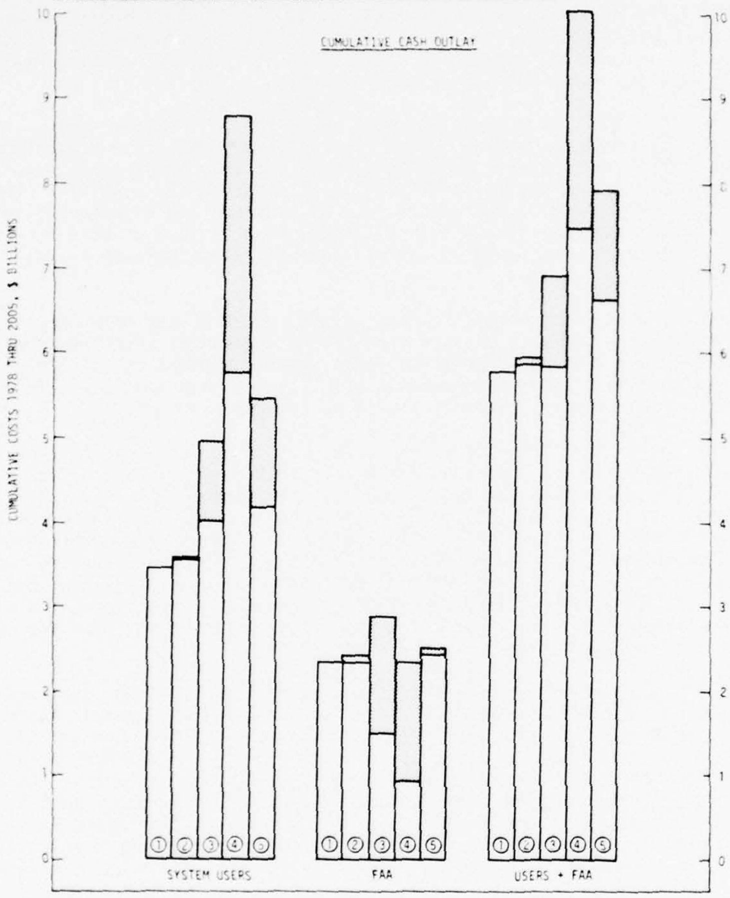


Figure 1. Air Navigation System Alternatives — Cost Comparison (Annual Inflation Rate = 7%; Cumulative Cash Outlay)

aviation. Thus the FAA was assessed the total costs required for the implementation and operation of the second-generation VORTAC system. The costs of non-civil aviation systems such as GPS were assumed to be borne by the designated operating agency, in this case the U.S. military, with the FAA incurring only those costs required to support and supplement that system as necessary to meet the U.S. civil aviation requirements. The shaded areas of Figure 1 reflect the range of costs attributable to variations from the nominal implementation and transition schedules noted in the legend.

The results shown are based on preliminary assumptions of user costs, i.e., avionics costs, and are not intended to be conclusive, particularly the cost effectiveness of satellite vs. non-satellite navigation; however, with the availability of satellite navigation, the potential expansion of Loran-C chains, and the upgrading of VORTAC, there is the potential of at least three different radio navigation systems which could satisfy civil air users in the U.S. The cost as shown for any one of these systems is about six to ten billion dollars (assuming variations in user costs and implementation schedules as shown). Clearly there is a need for careful planning, implementation, and potential phase out of radio navigation systems even when considering only the civil enroute air users in the U.S.

The complete quantification of radio navigation system life-cycle costs must include all classes of users for each system in order to provide complete system cost tradeoffs and subsequent implementation and/or phase-out decisions. The extrapolation of total life-cycle costs for all U.S. radio navigation

users (military and civil — air marine and land) for similar system scenarios as shown in Figure 1 clearly will result in enormous sums of money. Furthermore, the strong dependence of non-U.S.-users (military and civil) on U.S.-supported systems suggest the urgent need for international planning for the future of radio navigation.

The two essential steps in the navigation planning process are the quantification of user requirements and the comparisons of these requirements against present and future system capabilities. The status of user requirements and an assessment of current radio navigation system operational capabilities are summarized in the two succeeding sections.

3. RADIO NAVIGATION SYSTEM REQUIREMENTS

The first step in planning for an effective mix of radio navigation systems is a quantitative assessment of user requirements. Since the future of radio navigation systems will likely see the utilization of a small number of systems by a very large mix of air, sea, and land users both military and civil, the assessment of requirements must be achieved in a common format in order to compare candidate system capabilities against user requirements effectively. This has not been done to date across the different modal users and for both civil and military users. In fact, navigation "requirements" as defined today vary from international agreements/government regulations such as the requirement for VOR for IFR flight in the U.S. and ICAO states, to a simple statement of a particular user need such as the navigation accuracy required by a fishing boat to locate fishing grounds or traps.

Requirements vary from simple specifications of a navigation system for one vessel or aircraft (independent of other onboard systems) to complex system specifications where the navigation system requirement is but one part of the vehicle performance specification, e.g., weapon delivery, or of an overall traffic control system effectiveness and safety specification, e.g., that of air traffic control system. Also, whereas the most stringent functional requirements such as accuracy and coverage often influence the performance of a particular radio navigation system design, cost is also a valid user requirement; hence, when one system is utilized by a wide range of users, tradeoffs exist between overall system performance and user equipment costs.

The quantification of user requirements in some common format is not only complex because of the different requirements among users but also because many of the individual user requirements are changing. These changes include the number of users (which affects user equipment costs), the increasing use of radio navigation by new classes of users with unique requirements, and the increasing dependence of navigation user equipment on other onboard systems as well as external systems.

A significant increase in the number of civil radio navigation users can be expected for air, sea, and land applications. The projected increase in general aviation (including business aviation) and helicopter operations will change both the total number of air users as well as the mix (when compared to the relatively stable air carrier market). The projected increase in maritime world shipping, in the world fishing fleets, and other maritime applications as well as the more widespread use of radio navigation systems by maritime operators can be expected to result in a greatly increased number of maritime users. The use of radio navigation by land vehicles for position location is new, but could involve several million vehicles. Although military forces may not increase significantly over the next several decades, the wider spread use of radio navigation by the military can be anticipated. For example, the application of GPS to manpacks is a relatively new application.

In addition to the increased number of applications of radio navigation, the specific requirements of many users are changing. Requirements for off-shore, remote area, and oceanic navigation will increase coverage requirements for air navigation. The threat of marine collisions and groundings to the environment as well as human safety, and the existence of larger merchant ships with more difficult maneuvering requirements all tend to support the increased use of radio navigation systems by the maritime community as a replacement for and/or complement to audio-visual systems and dead-reckoning. The requirements of the military for precise weapon delivery, world-wide coverage, and all-weather operations are the most stringent of all which have lead to the development of GPS. The increases in fuel costs have lead to meaningful economic payoffs for more accurate track-keeping capability for both ships and aircraft.

The specifications of navigation requirements often are highly dependent on other onboard systems, external systems, and higher level performance measures. For example, requirements for military navigation systems for weapon delivery, track-keeping performance, and other mission requirements are coupled with the performance of guidance and control systems, and other avionics sensors. Civil air navigation requirements consider signal-in-space characteristics, onboard navigation receivers, flight technical and computer errors, as well as specified airspace requirements and safety (collision risk) measures. The desire of air carriers to improve energy management through profile optimization is dependent on accurate three-dimensional area navigation. Recent studies by the Coast Guard [2] have addressed the interaction of the mariner, the ship dynamics, the external environment (winds, current) and navigation system specification to achieve a safe and expeditious flow of traffic through waterways.

The variety, complexity and changing nature of user requirements, as well as the overall concern for the costs of operating duplicative radio navigation systems, have lead to several planning activities within the U.S. Government. The U.S. Department of Transportation (DOT), as the primary agency responsible for providing civil navigation systems, is conducting and sponsoring several radio navigation system planning and requirements activities. The DOT issues the U.S. National Plan for Navigation [3] which promulgates the current civil user requirements and system status. The equivalent military requirements are contained in the Joint Chiefs of Staff -- Master Navigation Plan.

3.1 Air Navigation Requirements

The requirements for international civil air navigation are established by agreement of the member states of the International Civil Aviation Organization (ICAO). U.S. domestic requirements include the broad requirements of the U.S. Department of Transportation National Plan for Navigation and more specific

requirements established by the Federal Aviation Administration Federal Aircraft Regulations (FAR's) and Advisory Circulars. Civil air navigation requirements are more complete than most other requirements in the sense that they consider interaction with other flight systems and they are generally specified based on an overall performance measure such as safety, minimum-distance economic considerations, and traffic flow/airspace planning considerations.

The world-wide air navigation requirements of ICAO member states which pertain to requirements specified within member states and for aircraft flying into these states are specified in five ICAO air navigation regional plan publications [4]. International Standards and recommended practices are specified in Annex 10 *Aeronautical Telecommunications* [5], which includes a definition of the short-range aid as being VHF omni-directional range (VOR) (required through 1985). Also included is a standard for Loran-A (since being discontinued), Consol or nondirectional beacons (NDB), and Distance Measuring Equipment (DME). More recently at the 9th Air Navigation Conference of ICAO, the Minimum Navigation Performance Specifications (MNPS) was adopted on a world-wide basis. This specification is intended to ensure safe separation of aircraft and at the same time enable operators to achieve maximum economic benefit from improvement in accuracy of navigation demonstrated in recent years. This concept will be implemented first in the North Atlantic Region (NAT) in order to establish standards subsequent to the withdrawal of Loran-A in December of 1977.

The MNPS specifies that the standard deviation of lateral track errors for the NAT region shall be less than 6.3 nm. Equivalently, an aircraft must stay within 12.6 nm of track for about 95% of the time. Also, the proportion of the total flight time spent by aircraft 30 nm or more off the cleared track shall be less than 5.3×10^{-4} (1 hour in about 200 flight hours), and the proportion of the total flight time spent by aircraft between 50 and 70 nm off the cleared track shall be less than 13×10^{-5} (1 hour in about 8,000 flight hours). Furthermore, such navigation performance capability shall be verified by the state of registry, the state of the operator (notice to Airmen, November 3, 1977). As an aid to operators for certifying navigation equipment, the FAA has issued Advisory Circular AC 120-33, "Operational Approval of Airborne Long-Range Navigation Systems for Flight Within the North Atlantic Minimum Navigation Performance Specification Airspace." The intent of these requirements is to reduce separation of aircraft in the NAT region from 120 nm to 60 nm laterally, and 2,000 feet vertically. Implementation of these standards generally can be achieved by using Omega or Omega/VLF combinations as an update to previously approved navigation systems, e.g., INS. Additional estimates of separation requirements (which may require better navigation systems than those which now exist) include for the NAT region:

Year	Lateral	Along Track
1977	30 nm	5 min
1990	20 nm	5 min
2000	15 nm	5 min

Similar standards are expected for other high density routes such as from the West Coast of the U.S. to the Pacific.

For the oceanic regions where dense traffic occurs, the criteria for requirements are not only safety but also lack of economic penalty from flying nonoptimum origin-to-destination tracks. (Hence the need to reduce separation requirements.) The requirements for long distance civil air navigation established by ICAO on 9 November 1965 include general categories listed in Table 1. Although these requirements again do not specify particular systems, they provide broad categories which must be complied with for any navigation system. The requirements for civil aviation in the U.S. are included in The Federal Aviation Regulations (FAR) part 91, 135, and 121. Also, Advisory Circular 90-45 specifies requirements for certifying area navigation systems.

The preceding paragraphs have summarized existing civil aviation regulatory requirements. These requirements are oriented toward current systems, international agreements, and oceanic navigation. The Federal Aviation Administration (FAA) has recently sponsored a study to define future requirements of U.S.

Table 1
General Considerations for Air Navigation Systems

(1)	That an aircraft can fly from origin to destination safely and economically.
(2)	That the needs of Air Traffic Service (ATS) and in some cases search and rescue (SAR) be satisfied.
(3)	A system be suitable for use in all aircraft types which require the same service considering such factors as size, weight, cost, power consumption, installation and certification requirements.
(4)	The level of performance may not be constant geographically or temporally because the position and track-keeping accuracy is a function of traffic density.
(5)	No one set of operational requirements can adequately reflect all the operating conditions encountered in all parts of the world.
(6)	The system must be capable of integration with the flight control system of the aircraft.

civil aviation which is independent of system capability [6]. This study focused on three specific geographic areas: (1) Continental U.S. (CONUS), (2) CONUS low-altitude off-shore, and (3) Alaska. User requirements in each geographic area were considered as well as current and anticipated future regulatory requirements.

The Instrument Flight Rules (IFR) navigation requirements for each of the three geographic areas are shown in Table 2 in terms of:

(1) Coverage
(2) Accuracy
(3) Operational Factors
(4) Capacity
(5) Compatibility
(6) Signal Reliability.

The format of these results is such that it can be combined with other user requirements, e.g., maritime, and ultimately aggregated with military requirements in order to determine cost vs. performance tradeoffs of candidate navigation systems.

Vertical coverage requirements are based on controlled airspace boundaries. Horizontal coverage requirements are based on current and projected air traffic needs.

Accuracy requirements are related to the route width associated with the route structure in the National Airspace System. The requirements are the same for CONUS, Alaska, and Offshore.

Operational factors relate to the navigation system interface with the other ATC components, namely, communications, surveillance, and safety. Flexibility, as used here, has to do with the ability to accommodate changes easily in route structure, including course, altitude and fixes or waypoints. Re-acquisition time relates to the time to activate the navigation system from an inactive state or to re-acquire the system following an interruption. This time is also representative of an upper bound for position-fix update rates in that the maximum time between updates is represented by the re-acquisition times. The rest of the operational factors are self-explanatory.

The capacity requirement is specified as being unlimited, which implies that the navigation system under consideration must be able to accommodate all aircraft accessing it at any time.

Compatibility refers to the interface between the navigation system under consideration and all other systems within the ATC system from an operational and electrical point of view.

Signal reliability is a specification of maximum signal outage time acceptable based on safety considerations.

3.2 Marine Navigation Requirements

No specific international requirements (government regulations) exist for marine navigation. A fundamental requirement based on the Safety of Life at Sea (SOLAS) convention [7] specifies basic onboard equipment. The primary objective for marine aids to navigation is similar (but less detailed) than the air navigation requirements, mainly, to support safe and economic movement of vessels from point of

Table 2
IFR Navigation System Requirements

USER	FLIGHT PHASE	REGION	COVERAGE: CONUS VERTICAL	COVERAGE: CONUS HORIZONTAL	COVERAGE: ALASKA VERTICAL	COVERAGE: ALASKA HORIZONTAL	COVERAGE: OFF-SHORE VERTICAL	COVERAGE: OFF-SHORE HORIZONTAL	ACCURACY (2D)	OPERATIONAL: FLEXIBILITY	OPERATIONAL: POSITION PRESENTATION	OPERATIONAL: COMMON INPUT FORMAT	OPERATIONAL: PILOT WORKLOAD	OPERATIONAL: FAILURE ALERTS	OPERATIONAL: POSITION RESOLUTION AMBIGUITY	OPERATIONAL: TIME TO RE-ACQUIRE	CAPACITY	COMPATIBILITY	SIGNAL RELIABILITY
IFR	ENROUTE	MOUNTAINOUS	2000 AGL TO FL 600	TOTAL	2000 AGL TO FL 600	TOTAL	500 AGL TO 10,000 MSL	200 NMI OFF-SHORE	±4 NMI 95%	YES(5)	COURSE DEVIATION, NMI (6) LAT/LON	YES	(7)	MUST BE AVAILABLE	LESS THAN 0.5% OF THE TIME	1-2 MIN	UNLIMITED	AS PER PREVAILING SPECS	FAIL-SOFT
IFR	ENROUTE	NON-MOUNTAINOUS	2000 AGL TO FL 600	(1)	2000 AGL TO FL 600	TOTAL	NOT APPLICABLE	NOT APPLICABLE	±4 NMI 95%	YES	COURSE DEVIATION	YES	(7)	MUST BE AVAILABLE	PRECLUDE VIA SYSTEM DESIGN	0.5-1 MIN	UNLIMITED	AS PER PREVAILING SPECS	FAIL-SOFT
IFR	TERMINAL	HIGH DENSITY	200 AGL TO 14,500 MSL	(2)	200 AGL TO 14,500 MSL	(2)	200 AGL TO 10,000 MSL	(2)	(4,9) ±2 NMI 95%	YES	COURSE DEVIATION	YES	(8)	MUST BE AVAILABLE	PRECLUDE VIA SYSTEM DESIGN	0.25-0.5 MIN	UNLIMITED	AS PER PREVAILING SPECS	FAIL-SOFT
IFR	TERMINAL	LOW DENSITY	200 AGL TO 14,500 MSL	(2)	200 AGL TO 14,500 MSL	(2)	200 AGL TO 10,000 MSL	(2)	(9) ±4 NMI 95%	YES	COURSE DEVIATION	YES	(8)	MUST BE AVAILABLE	PRECLUDE VIA SYSTEM DESIGN	0.5-1 MIN	UNLIMITED	AS PER PREVAILING SPECS	FAIL-SOFT
IFR	NON-PRECISION APPROACH		250 AGL TO 14,500 MSL	(3)	250 AGL TO 14,500 MSL	(3)	250 AGL TO 10,000 MSL	(3)	±1.5 NMI	YES	COURSE DEVIATION	YES	(8)	MUST BE AVAILABLE	PRECLUDE VIA SYSTEM DESIGN	0.25 MIN	UNLIMITED	AS PER PREVAILING SPECS	FAIL-SOFT

(1) Equivalent to current and increasing with time to reflect projected traffic density increases.
(2) All terminal areas being served currently and those projected to be serviced.
(3) All airports currently with non-precision approach procedures and those where such procedures are expected to be required.
(4) 3D - ±2.0 NMI, 4D - ±.15 NMI
(5) Not a hard requirement, but does have significant cost impact.
(6) Enroute preplanned direct only.
(7) Less than or equal to single waypoint VORTAC-based RNAV system.
(8) Less than or equal to dual waypoint VORTAC-based RNAV system.
(9) ±2 NMI in terminal maneuvering area (within 15 NMI of airport) ±4 NMI beyond 15 NMI from the airport.

departure to point of arrival. Adequate position-fixing capability is required which is not limited by geographic location or the environment (wind, sea state, visibility).

The same general navigation considerations included in Table 1 also apply to marine navigation. Specific needs of marine navigation have been promulgated in the National Plan for Navigation for: (1) high seas, (2) coastal confluence areas, and (3) harbor/harbor entrances. These requirements are summarized in Table 3.

Table 3
Marine Navigation Requirements

ENVIRONMENT	GENERAL REQUIREMENTS	AVAILABILITY	COVERAGE	ACCURACY
HIGH SEAS				
• POSITION FIXING BY VISUAL REFERENCE TO LAND OR OTHER FIXED OR FLOATING AIDS IS NOT PRACTICAL	• POSITION FIXES REQUIRED TO EFFICIENTLY DIRECT A COURSE TOWARD DESIRED DESTINATION • OCEANIC AND SURVEY ENDEAVORS REQUIRE POSITION FIXING RELATIVE TO A GEODESY • MUST BE ADEQUATE TO AVOID CHARTED OBSTRUCTIONS INCLUDING REEFS, REMOTE ISLANDS AND COASTS WHERE OTHER NAVAIDS DO NOT EXIST.	• CONTINUOUS SYSTEM DESIRABLE • UPPER LIMIT OF 2 HOURS	• GENERALLY WORLDWIDE • U.S. COMMERCIAL MARITIME INTERESTS 90% SATISFIED BY COMPLETE COVERAGE OF NORTH ATLANTIC AND NORTH PACIFIC	• 4 NMI 95% OF THE TIME • FUTURE GOAL OF 2 NMI 95% OF THE TIME BY YEAR 2000
COASTAL CONFLUENCE				
• WATERS CONTIGUOUS TO MAJOR LAND MASSES OR ISLAND GROUPS WHERE TRANSOCEANIC TRAFFIC PATTERNS CONVERGE IN APPROACHING DESTINATION AREAS. • INTERPORT TRAFFIC EXISTS IN PATTERNS PARALLELING COAST LINES • PRESENCE OF SHALLOW WATER • INCREASED CONGESTION • VARIETY OF VESSEL TYPES • POINT-TO-POINT COAST-WISE TRADE • DIRECTED VESSEL MOVEMENT (FISHING OR SURVEY VESSELS) • RECREATIONAL CRAFT • AUDIO/VISUAL AIDS MAY OR MAY NOT EXIST	• POSITION FIXING CAPABILITY WHERE PROXIMITY OF LAND, USER MIX AND TRAFFIC DENSITY IMPOSE HIGH DEGREE OF HAZARD FOR COLLISION OR STRANDING • SHOULD PROVIDE CAPABILITY TO SUPPORT ANTI-COLLISION PROCEDURES SUCH AS TRAFFIC SEPARATION SCHEMES • SYSTEM SHOULD RESULT IN FEWER DELAYS DUE TO REDUCED VISIBILITY AND PROVIDE CAPABILITY TO ACCURATELY ACQUIRE A LANDFALL • SPECIALIZED SEGMENTS (FOR REPEATABILITY) MAY EXIST, E.G., FISHERMAN FOLLOWING PARALLEL TRACKS OR OPERATING CLOSE TO KNOWN BOTTOM OBSTRUCTIONS	• CONTINUOUS INFORMATION WITH HIGH RELIABILITY • MUST NOT BE LIMITED BY VISIBILITY OR LINE-OF-SIGHT DISTANCES	• NO SPECIFIC DEFINITION; MAY VARY FROM 21 NMI TO 120 NMI	• GENERAL GOAL OF POSITION FIXING TO A REPEATABLE RMS ACCURACY OF 1/4 NMI UP TO 50 NMI OFFSHORE WITH REDUCED ACCURACY AVAILABLE AT GREATER DISTANCES • SERVICE OF USEFUL QUALITY SUCH THAT NO VESSEL IS LOST OR ENDANGERED BECAUSE OF LACK OF POSITION KNOWLEDGE WITHIN THE ECONOMIC REACH OF ALL MARINERS
HARBOR AND HARBOR ENTRANCE				
• WATERS INSIDE THE MOUTHS OF RIVERS AND BAYS INCLUDING INLAND LAKES WHERE TERMINAL FACILITIES ARE LOCATED • OPERATIONS IN PROXIMITY TO SHALLOW WATER, NARROW CHANNELS AND ENTRANCES, NEAR OBSTRUCTIONS, NEAR OTHER VESSELS • REQUIRES HIGHEST DEGREE OF MANEUVERING PRECISION AND VESSEL CONTROL TO AVOID DANGERS OF COLLISION OR STRANDING • AUDIO VISUAL/AIDS EXIST • EFFECTS OF TIDES AND CURRENT MORE IMPORTANT	• NO HARD REQUIREMENTS FOR RADIO NAVIGATION AIDS IN U.S. HARBOR/HARBOR ENTRANCES • SOME REQUIREMENTS FOR RADIO NAVIGATION AIDS IN CERTAIN FOREIGN PORTS (ROTTERDAM, GOTTENBURG) • GENERAL REQUIREMENT TO COMBINE POSITION FIXING (AND IN SOME CASES RATE INFORMATION, E.G., MANEUVERING VLCC'S IN RESTRICTED WATERS) WITH PROCEDURAL INTELLIGENCE	• CONTINUOUSLY AVAILABLE WITH HIGH RELIABILITY	• ONLY IMPLEMENTED WHERE SAFETY AND ECONOMIC PENALTIES IMPOSED BY CURRENT LIMITATIONS ARE DETERMINED TO BE SUFFICIENT TO SUPPORT THE COSTS	• SUFFICIENT TO PROCEED AS IF THE VISUAL AID DID NOT EXIST

The requirements for marine navigation in the high-seas environment as shown in Table 3 are in general similar to those for air navigation. The major exceptions are data rate and the method of display. The similarity between air and marine long-distance navigation aid requirements would appear to provide motivation for international agreements in order to prevent costly alternatives of redundant systems.

The need for establishing specific requirements for coastal confluence areas can be expected to continue because of the consequences of collisions, groundings and strandings. Recent proposed rules by the Coast Guard include for U.S. coastal waters a requirement for all vessels over 1600 gross tons to be equipped with one of the following systems:

(1) Loran-C

(2) Satellite and some dead-reckoning device such as a Doppler Sonar

(3) Any system equivalent to Loran-C if approved by the U.S. Coast Guard Commandant.

Specific performance standards for Loran-C have been established by IMCO [8].

The trend toward wider use of radio navigation aids for harbor/harbor entrances is also apparent. This trend is clearly motivated by larger deep-draft ships, more dense traffic, the impact of a collision between ships containing hazardous cargoes, and the cost of maintaining audio-visual aids. The current requirement for Decca in Rotterdam (the brown box carried aboard by the pilots), and the expected requirements for a radio navigation aid in Gothenburg are good examples of this trend. Also, experience of maneuvering large ships in restricted areas (such as VLCC's maneuvering in Antifer Harbor) has indicated the need for radio navigation systems with sufficient accuracy to provide rate information. This information cannot generally be provided with sufficient accuracy or speed by visual perceptions alone.

3.3 Ground Vehicle Requirements

There are presently no firm requirements for ground vehicle navigation. The large number of potential ground vehicles which may require radio navigation have been identified by Gilbert in a subsequent chapter. Several demonstration projects have been sponsored by the U.S. Government to evaluate the use of Loran-C by ground vehicles such as truck fleets, police and emergency users. Although neither firm requirements nor widespread operational usage of radio navigation currently exists, the potentially large number of users in this group is certainly sufficient to consider this application in any future navigation system planning.

4. RADIO NAVIGATION SYSTEMS — CURRENT STATUS

This section summarizes the current status of major radio navigation systems which are in use at this time. Included are a summary of the technical and operational characteristics, the current status (availability and coverage), and current utilization. Other summaries of navigation system operation and performance and usage are included in Refs. 9-36. Emphasis here is placed on systems which are relatively new, are not completed or are being discontinued or expanded. Also, those systems which are utilized extensively on an international basis or are currently standards for usage in certain countries are also discussed in detail.

The U.S. policy on radio navigation aids has been clarified somewhat in recent years by supporting the use of certain systems in specific areas for particular modes. The U.S. Department of Transportation policy on radio navigation aids is defined by the National Plan for Navigation [3]. This policy includes:

(1) use of VORTAC as the standard short-range navigation aid for civil aviation;

(2) the use of Loran-C as the government-funded navigation aid for coastal confluence marine navigation;

(3) the continuance of support for nondirectional beacons for both air and marine navigation. Particularly, as a backup for civil air navigation, a major part of the navigation system where VORTAC coverage does not exist such as in Alaska, and as a low-cost system for both air and marine users;

(4) the planned phase-out of Loran-A;

(5) the potential use of NAVSTAR/GPS for civil applications.

In addition, the U.S. government supports the international Omega hyperbolic radio navigation system. Military systems include TACAN for short-range air navigation, Omega for air and marine long-range navigation and submarine navigation, Loran-C and -D for special purpose use such as weapon delivery, and the transit satellite system for long-range marine navigation systems.

The above systems are also used extensively on an international basis. In many non-U.S. coastal confluence areas and harbor entrances, Decca is used instead of Loran-C. The transit system is described in a succeeding paper and is therefore not discussed here. Also, other radio navigation systems such as JTID's, GPS and PLRS are covered in succeeding papers. Specialized systems such as precision approach radars, instrument landing systems, Consul, RACONS, are not discussed here because of their limited usage or as terminal guidance systems.

The following systems are summarized in succeeding paragraphs:

- Loran-A
- Loran-C
- Omega
- Decca
- VOR/DME
- TACAN

Figure 2 summarizes the implementation/operational schedule for these systems [11].

4.1 Loran-A

Description

Loran-A is a High Frequency (HF) hyperbolic radio navigation system developed during World War II. Loran-A chains, usually installed along a coastline, are normally comprised of a master and two secondary stations. The baseline length between a master and one of the secondary stations is generally on the order of 200 miles. At these frequencies, attenuations over land are high, but reception range over seawater extends to 800 miles. The transmitted signal is a pulsed carrier. Station pairs are identified by frequency and pulse-repitition rates. Each chain uses one of three carrier frequencies; 1850, 1900 or 1950 kHz. Twenty-four different pulse repetition rates are used ranging from 20 pps to 34-1/9 pps. Time difference measurements are made by observing the time delay between the master pulse and the corresponding secondary pulse on an envelope basis by using an oscilloscope.

Accuracy

Absolute — 1 to 5 nm (RMS)

Repeatability — Approaches 1,000 feet (RMS).

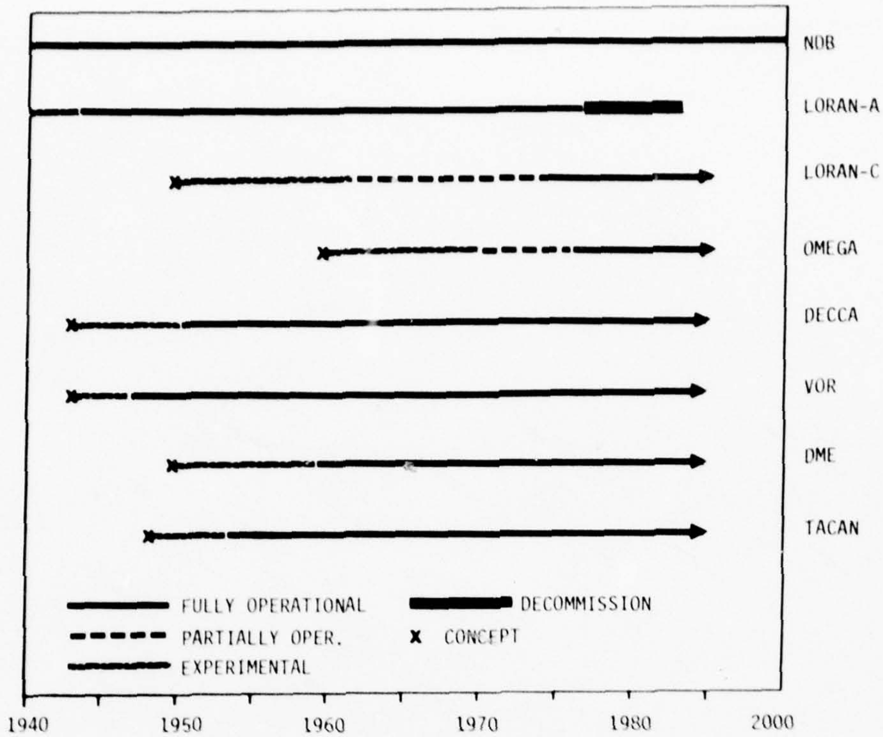


Figure 2. Implementation/Operational Schedule

Coverage

The Loran-A coverage diagram is shown in Figure 3. Ground-wave coverage is provided over most of the navigable coastal areas of the Northern Hemisphere. Nighttime coverage is available over most of the oceanic routes of the Northern Hemisphere.

Status

By the end of the World War II, over 70 stations were in operation. Requirements by the U.S. Military, merchant shipping and transoceanic airlines have caused the system to remain operational to date. It has been estimated that there are between 50,000 and 100,000 users of Loran-A today. This includes a large number of small fishing craft whose population is difficult to assess because their Loran-A sets were obtained as World War II surplus gear or second hand. Loran-A was one of the systems under consideration by the U.S. Coast Guard to meet the Coastal Confluence Region requirements. After exhaustive studies and tradeoffs, it was decided to select Loran-C to serve the Coastal Confluence Region. This was officially announced in May 1975. At the same time, an orderly phase-out schedule of the Loran-A network was announced. The planned termination dates for the U.S.-operated Loran-A chains are:

DOMESTIC

Aleutian Islands	31 July 1979
Gulf of Alaska	31 December 1979
Hawaiian Islands	31 July 1979
West Coast	31 December 1979
Caribbean	31 December 1980
East Coast	31 December 1980
Gulf of Mexico	31 December 1980

All U.S.-operated overseas stations have been phased out. Some additional NATO stations are still operational.

The primary replacements for Loran-A are expected to be:

(1) Omega for oceanic air navigation and marine.
(2) Loran-C for coastal confluence and some non-U.S. areas.
(3) Transit satellites for marine high-sea navigation.

4.2 Loran-C

Description

Loran-C is a low-frequency (LF) hyperbolic radio navigation system developed by the Department of Defense during the 1950's to meet operational military requirements. The first Loran-C chain, located along the U.S. East Coast became operational during 1959-1960. Today there are nine chains operational, with a total of twelve expected by 1980. The transmitted signal is a pulsed 100 kHz carrier. Compared

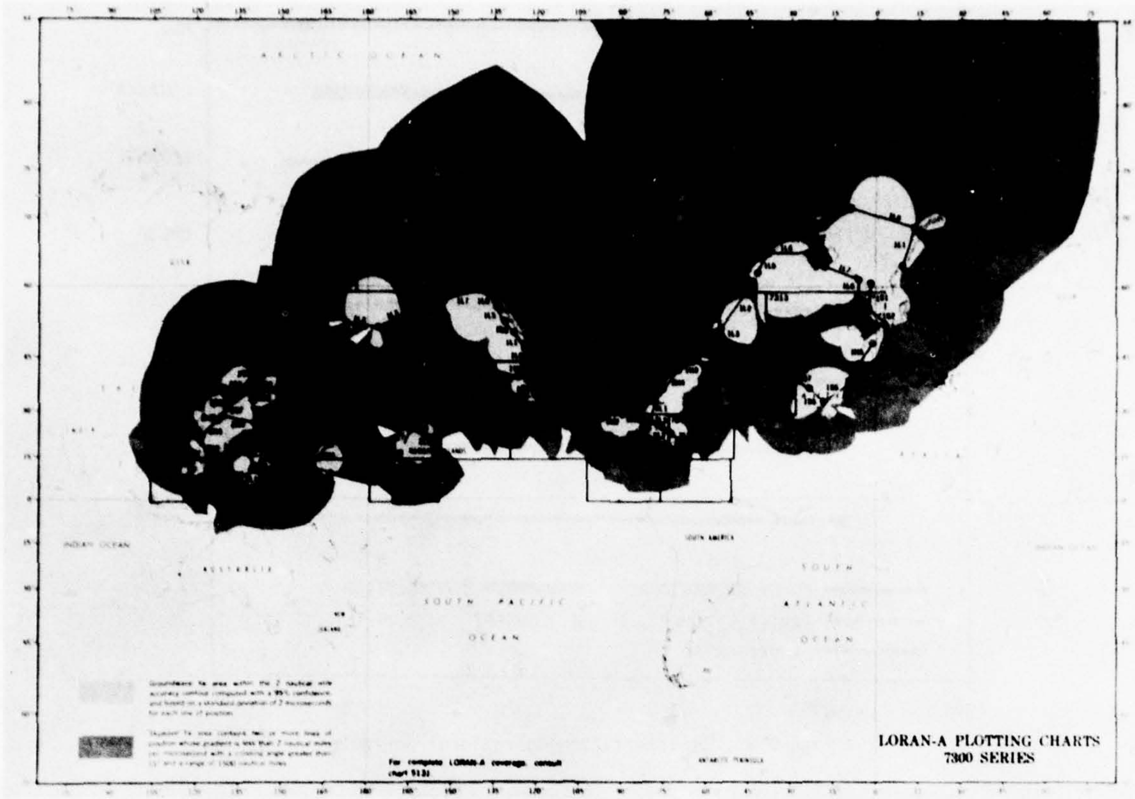


Figure 3. Loran-A Coverage Diagram

to the HF Loran-A frequencies, propagation losses, particularly over land, are lower, and propagation stability is higher at 100 kHz.

These factors permit the use of longer baselines between stations which increases the useful coverage area for a given number of stations. Baseline lengths are on the order of 600 miles to 800 miles. Useful signals can be received at ranges of 800 to 1200 miles from the stations. Loran-C chains are uniquely identified by their pulse Group Repetition Interval (GRI). GRI's are assigned between 40,000 and 99,990 microseconds in 10-microsecond steps. During one GRI, the master transmits 9 pulses and each secondary transmits 8 pulses as illustrated in Figure 4. The ninth pulse from the master, spaced 2000 microseconds after the eighth, is intended to aid in identifying the master signal. Each secondary station is assigned a unique coding delay (CD) which is used for identifying the secondary signals within a chain. The coding delay is the time between the master signal and the secondary signal as observed at the secondary station. All Loran-C transmitting stations are equipped with cessium frequency standards which allow the systems to serve as a very stable frequency reference for any user within ground-wave range of any Loran-C stations. Similarly, precise time is made available by synchronizing the transmissions to Universal Coordinated Time (UTC) as determined by the U.S. Naval Observatory.

Accuracy

Absolute — 1/4 - 1 nm (RMS)

Repeatability — 300 ft (RMS).

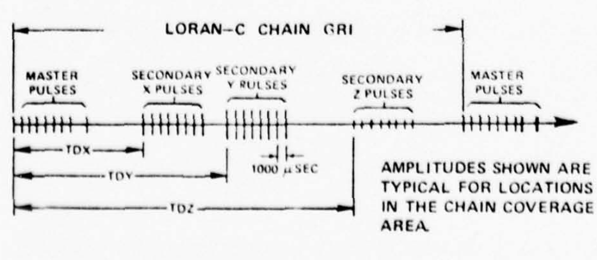


Figure 4. Example of Received Loran-C Signal

Coverage

The Loran-C coverage diagram is shown in Figure 5. Ground-wave coverage is provided over most of the North Atlantic, the Mediterranean, the Coastal Confluence Regions of CONUS, Alaska, and Hawaii. The Western Pacific is covered between Japan, the Philippines, and Guam. The Southeast Asia chain is now decommissioned. In addition, Alaska is covered except for the north slope and CONUS is covered except for the region between the Rocky Mountains and the Mississippi River.

There are also privately owned chains using low-power commercial transmitters operating in the Gulf of Mexico, the Java Sea, and the North Sea [35]. The U.S. Coast Guard is operating a low power chain covering the St. Mary's River [35]. The U.S. Air Force operates Loran-D chains covering the southwestern southeastern, U.S. test ranges and Central European test ranges [35]. Loran-D is a military tactical system which operates on the same frequency as Loran-C, but with lower power and a higher pulse rate. Military receivers are designed to operate with either Loran-C or Loran-D signals.

Status

The following chains are currently in operation by the U.S.:

Coast Guard:

1. U.S. East Coast Chain - Rate 9930
2. U.S. West Coast Chain - Rate 9940
3. West Canadian Chain - Rate 5990
4. North Pacific Chain - Rate 9990
5. Gulf of Alaska Chain - Rate 7960
6. Northwest Pacific Chain - Rate 9970
7. Central Pacific Chain - Rate 4990
8. North Atlantic Chain - Rate 7930
9. Norwegian Sea Chain - Rate 7970
10. Mediterranean Sea Chain - Rate 7990
11. St. Mary's River Chain - Rate 4970, Low Power

U.S. Air Force

1. Southeast U.S. Test Chain - Rate 3970, Loran-C/D
2. Utah Test Chain - Rate 4970, Loran-D
3. Central European Chain - Rate 2970, Loran-D

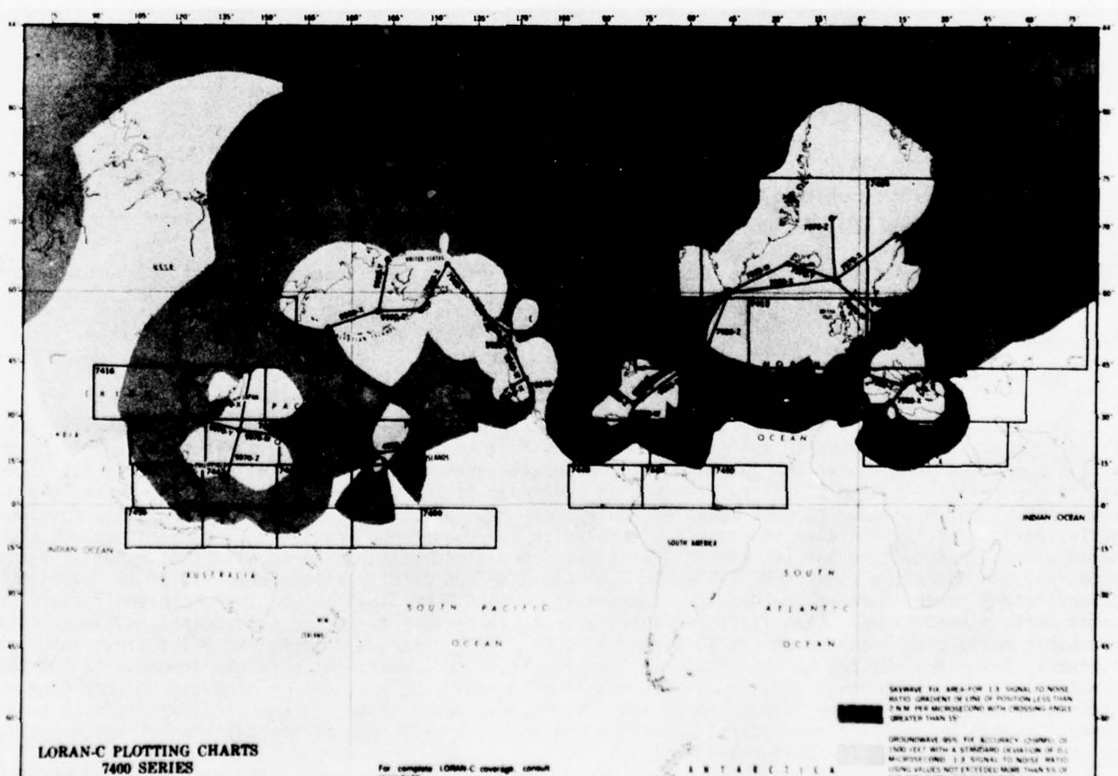


Figure 5. Loran-C Coverage Diagram

U.S. Army

1. Fort Hood Chain - Rate 4970, Loran-D

Private Industry

1. Gulf of Mexico - Rate 4864, Low Power, Industrial Location Service
2. Eastern USSR Chain - Rate 5000
3. Western USSR Chain - Rate 8000

The following additional chains are scheduled to become operational by the U.S. Coast Guard:

1. Gulf of Mexico Chain (Operational 7/78)
2. Great Lakes Chain - Rate 9960 (Operational 2/80)
3. Mid-Continental U.S. (funding not approved).

Most of the present Loran-C users are either military or maritime. Military applications include aircraft, ships, manpacks, and some land vehicles. Loran-C has been in use on the U.S. Navy's Fleet Ballistic Missile Submarines since about 1960. The U.S. Air Force has been using Loran-C on aircraft for tactical applications since 1968 and has used Loran-D as a tactical ground system.

The U.S. Coast Guard randomly sampled navigation practice and equipment of the maritime fleet in mid-1975 [37]. Of the vessels boarded, 40% were equipped with Loran-C receivers. Vessels under U.S. Flag comprised 21% of the sample and 70% of these were equipped with Loran-C.

Loran-C is one of the options available to the users of Loran-A, but to date there are not any serviceable low-cost receivers available for civil aircraft.

Previous proposed rule-making notices by the U.S. Coast Guard have established requirements for Loran-C for all vessels of 1600 gross tons or more. Recent modifications to these proposed rules are expected to require Loran-C or an equivalent. It is estimated that to satisfy this requirement, approximately 400 U.S. to 4000 foreign-flag vessels would have to equip with Loran-C or an equivalent system.

Loran-C is also being considered as a position locating system for civil emergency vehicles, railroad cars, and trucks. These considerations have been stimulated by the Department of Transportation studies and experimental evaluations. These systems are often referred to as automatic vehicle monitoring (AVM) equipment. Developments in this area and future implementation is expected to provide a large new market for radio navigation aids.

The FAA is evaluating Loran-C as a replacement or supplement to the VOR/DME air navigation system over CONUS and Alaska. The primary advantage offered by Loran-C is total all-altitude coverage down to ground level which is not possible with VOR/DME. The drawback of this system is with large service area, and potentially large number of aircraft affected by the outage of a single transmitting station. Other major advantages include area navigation capability and coverage in remote and off-shore areas. Other major disadvantages include a lack of airborne operational experience, airborne receivers and complete CONUS coverage.

Advantages offered by Loran-C for air navigation include:

(1) the area navigation capability in all areas;

(2) coverage in remote and off-shore areas which is not available from VORTAC; and

(3) coverage down to ground level.

Disadvantages for air navigation include the large impact of a single transmitting station outage, lack of operational experience, availability of airborne receivers, and the present lack of CONUS coverage.

4.3 Omega

Description

Omega is a very low frequency (VLF) radio navigation system operating in the internationally allocated frequency band between 10 and 14 kHz. At these frequencies, the earth's surface and the ionosphere act as a wave guide which allows the signals to propagate over long distances with relatively low attenuation and relatively high stability. A further advantage of these very low frequencies is the relatively low attenuation in passing through seawater, only about 1.0 dB/ft, thus providing a capability for submarine navigation. The system is capable of providing all-weather navigational service throughout the world with a transmitting complex of eight stations. The eight stations along with their identification, location, and operating agency are listed in Table 4. The permanent stations transmit at 10 KW which is sufficient power at these frequencies to propagate a signal half-way around the world and farther under certain conditions. Each station transmits in a synchronous format as illustrated in Figure 6. The basic navigation frequencies are 10.2 kHz, 13.6 kHz, 11-1/3 kHz, and 11.05 kHz. The first three are currently being transmitted by all stations. Modifications to the transmitters are underway to add the capability to transmit at 11.05 kHz. The reasons for transmitting more than a single frequency from each station is to provide a lane resolution capability. However, additional frequencies also tend to enhance overall system reliability since the attenuation and modal interference of the different frequencies can be quite different under certain conditions.

The need for lane resolution is illustrated by Figure 7. The basis for determining position within the Omega system is by phase measurement. Either of two techniques can be used — direct phase measurement of the signals from each station used (which requires a very stable clock synchronized to the Omega transmissions), or phase difference measurements of the signals from station pairs. The result of the

Table 4
Omega Transmitting Station Network

STANDARD LETTER DESIGNATION	LOCATION	COGNIZANT AGENCY
A	ALDRA, NORWAY	NORWEGIAN TELECOMMUNICATIONS ADMINISTRATION
B 1/	MONROVIA, LIBERIA	LIBERIAN DEPARTMENT OF COMMERCE, INDUSTRY AND TRANSPORTATION
C	HAIKU, HAWAII	U.S. COAST GUARD
D	LaMOURE, NORTH DAKOTA	U.S. COAST GUARD
E	LaREUNION	FRENCH NAVY
F	GOLFO NUEVO, ARGENTINA	ARGENTINE NAVY
G 2/	AUSTRALIA (TENTATIVE)	AUSTRALIAN DEPARTMENT OF TRANSPORTATION
H	TSUSHIMA, JAPAN	JAPANESE MARITIME SAFETY AGENCY

1/ STATION B, LIBERIA, IS OPERATED BY A U.S. CONTRACTOR SPONSORED BY THE U.S. GOVERNMENT

2/ A TEMPORARY STATION AT TRINIDAD (10° 42'N/61°38'W), OPERATED BY THE U.S. COAST GUARD IS TRANSMITTING IN THE G TIME SLOT. THE TRINIDAD STATION WILL CONTINUE OPERATION UNTIL FURTHER NOTICE.

STATION \ SEGMENT	A	B	C	D	E	F	G	H
NORWAY (A)	10.2	13.6	11⅓	12.1	12.1	11.05	12.1	12.1
LIBERIA (B)	12.0	10.2	13.6	11⅓	12.0	12.0	11.05	12.0
HAWAII (C)	11.8	11.8	10.2	13.6	11⅓	11.8	11.8	11.05
N. DAKOTA (D)	11.05	13.1	13.1	10.2	13.6	11⅓	13.1	13.1
LA REUNION (E)	f_1	11.05	f_1	f_1	10.2	13.6	11⅓	f_1
ARGENTINA (F)	12.9	12.9	11.05	12.9	12.9	10.2	13.6	11⅓
SOUTH PACIFIC (G)* AREA	11⅓	f_1	f_1	11.05	f_1	f_1	10.2	13.6
JAPAN (H)	13.6	11⅓	f_1	f_1	11.05	f_1	f_1	10.2
	0.9	1.0	1.1	1.2	1.1	0.9	1.2	1.0

0.0 1.1 2.3 3.6 5.0 6.3 7.4 8.8 10.0

TIME (SECONDS)

* TRINIDAD TEMPORARILY FILLING G SLOT

POLICY FOR UNIQUE FREQUENCY TRANSMISSION IN UNUSED FORMAT SEGMENTS BEING FINALIZED

PROPOSED FULL FORMAT IS SHOWN:
- f_1 IS UNIQUE FREQUENCY AT EACH STATION
- 11.05 IS FOURTH NAVIGATION FREQUENCY

Figure 6. Omega Signal Transmission Format

first method is a family of circular lines-of-position (LOP's) centered about the station and spaced one wavelength apart. The second technique yields a family of hyperbolic lines-of-position (LOP's), spaced one-half wavelength apart. As can be seen, there are many LOP's valid for any given measurement. Automatic Omega receivers are designed to keep track of the correct lane by counting so that the distance traveled is related to a whole number of lanes plus a fraction, the lane fraction being directly related to the phase measurement.

Given that the Omega receiver is initially set to the correct lane at the beginning of travel and that continuous, correct phase measurements are made during travel, correct position will be known within the accuracy of the Omega system; therefore, correct initialization or re-initialization of the receiver requires that position be known to one-half of a lane width. The hyperbolic lane width at 10.2 kHz along the baseline is about 8 miles. Clearly, wider lanes are desirable. Wider lanes can be achieved by the

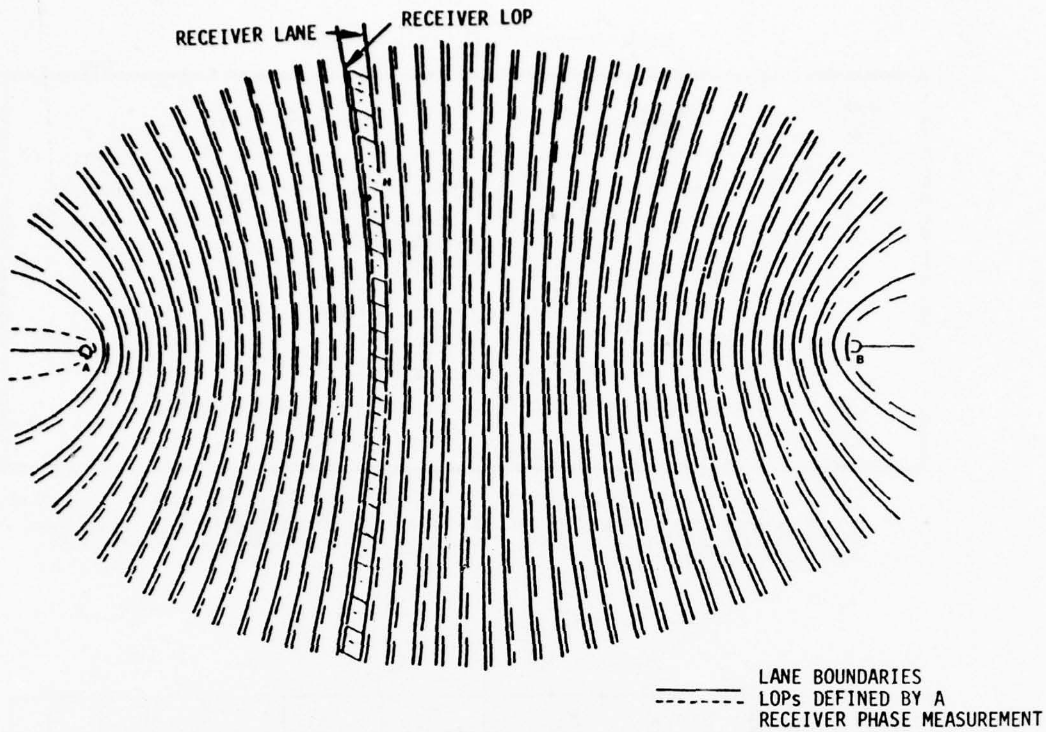


Figure 7. Family of Hyperbolic Lanes

combined use of the additional signals as illustrated in Figure 8. A 3.4 kHz phase measurement is used to resolve the 10.2 kHz phase position to point A or B. A phase measurement at 1.1-1/3 kHz resolves the 3.4 kHz ambiguity to point A. The 3.4 kHz phase measurement is obtained by taking the difference between the 10.2 kHz and 13.6 kHz phase measurements. Similarly, the difference between the 10.2 kHz and 11-1/3 kHz phase measurements yields a 1.1-1/3 kHz phase measuremement. The hyperbolic lane widths along the baseline for 3.4 kHz and 1.1-1/3 kHz are 24 and 72 miles, respectively. Wider lanes can be achieved in a similar manner.

Automatic receivers achieve station identification during initialization by correlating the received signals with an internally generated format. The addition of unique frequencies, as indicated in Figure 6,

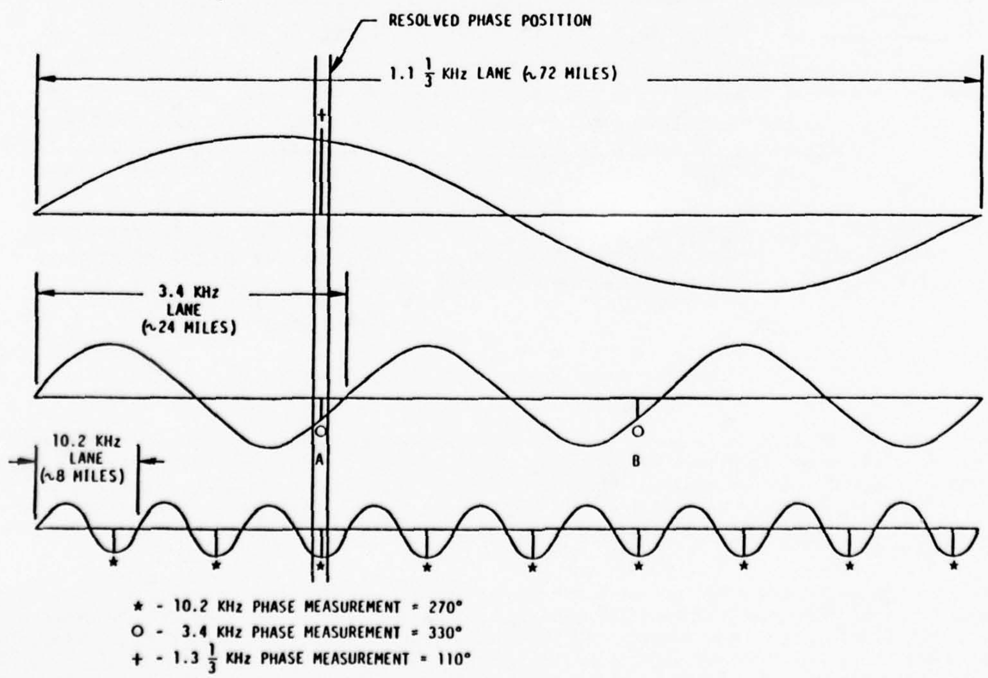


Figure 8. Multiple Frequency Lane Resolution

will permit simpler station identification. In addition, the unique frequencies will tend to enhance overall system reliability by doubling the duty cycle of usable signals from each station.

An important element in the use of Omega is the need to correct for non-uniform propagation velocity. Due to the non-uniform electromagnetic properties of the earth's surface and the ionosphere together with diurnal variations and the presence of a geomagnetic field, the wave fronts are appreciably non-circular. Thus, the phase velocity of the wave depends significantly on the direction of propagation; and consequently, a user's position, as determined by received phase and charts alone, will not coincide with his true position. Much experimental and theoretical investigation of VLF wave propagation has contributed to the calculation of these phase discrepancies. The results of these studies have been synthesized into a computer program which generates tables of phase corrections, termed propagation corrections (PPC's). Individual corrections are tabulated for 4° x 4° regions of the world and a navigator simply algebraically adds the PPC difference computed for his region to the measured phase difference to obtain a more accurate LOP. These tables have undergone continual refinement and, currently, use of PPC's can yield positional accuracies of approximately one nautical mile [38].

A typical example propagation correction function for a given user location, transmitter, frequency, and day is shown in Figure 9. These corrections (PPC's) are made available to users in the form of tables or computer programs.

In order to achieve satisfactory operation, a great deal of care is required in the installation of the Omega receiver and its associated antenna. In order to achieve world-wide coverage at all times of the year, the receiving system is required to operate down to signal-to-noise ratios of -20 dB (100 Hz BW). In addition to the normal background noise level, there are potential interference sources on the vehicle. Some of the higher harmonics of 60 Hz and 400 Hz are on or very near the Omega frequencies. In addition, precipitation static (P-static) can cause severe interference problems. Recent engineering developments in the antenna design and location as well as to the receiver installation have identified solutions to the installation problems on board aircraft. Acceptable approaches have been defined by RTCA and AEEC working groups; however, more attention will be required before the installation problems are well understood and resolved for marine application.

Accuracy

Absolute — 1 to 2 nm (RMS)

Differential Mode — 1/4 to 1 nm (RMS).

Coverage

A simplified average composite signal coverage map is illustrated in Figure 10. The contours are based on predictions and may tend to be somewhat optimistic. Coverage at noon mid-summer is reduced and some studies indicate larger areas of only three-station coverage than those shown in Figure 10. Areas with only three-station coverage provide no redundancy for hyperbolic operation in the event one of the three stations is off the air. It should be noted that the predicted coverage contours shown in Figure 10 are based on the full complement of eight stations including the future Station G expected to be operational in Australia by 1980.

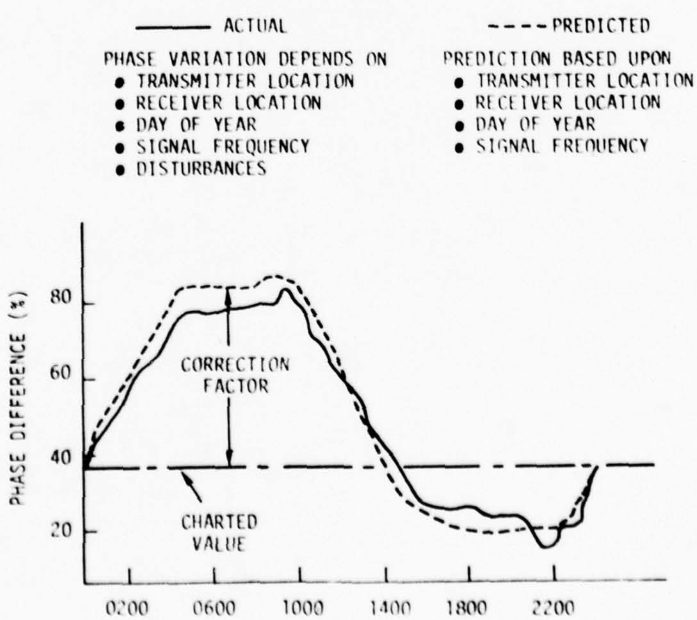
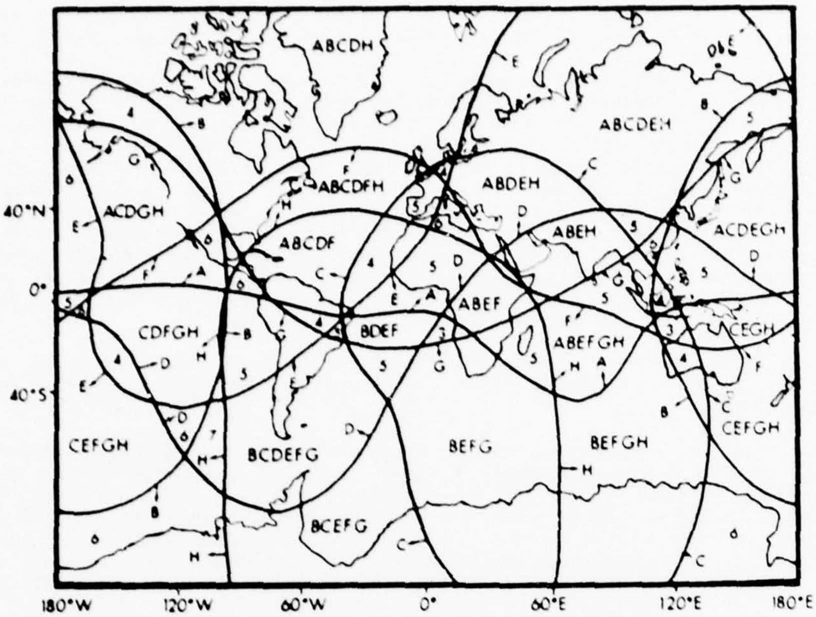


Figure 9. Typical Phase Variation Curve.



Letters indicate coverage by designated transmitting station
Numerals indicate number of transmitting station signals available in that region.

Contours are drawn for S/N = −20 db (100 Hz BW).
S/N is computed for Mar–Apr–May quarter and averaged over 24 hours.

Figure 10. Composite Signal Coverage Map

Status

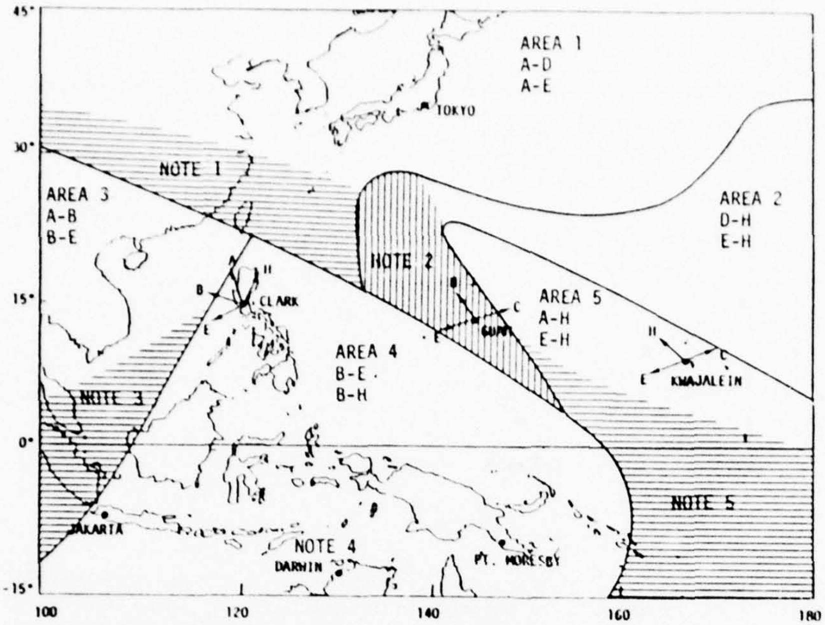
Omega is an international system dedicated to providing a global all-weather navigation and positioning capability of moderate accuracy. The Omega Navigation System Operations Detail (ONSOD) of the U.S. Coast Guard is the responsible agency for the United States. ONSOD oversees U.S. interests in Omega, operates two permanent stations and is conducting a signal-monitoring and data-update program. As part of their monitoring program, ONSOD issues a weekly status report on Omega system operation. The report is available to users by mail or TWX. An abbreviated version is also available on a recorded message via the telephone. ONSOD is also operating the temporary station at Trinidad. There is speculation that this station will remain on the air until the eighth permanent station becomes operational. In addition, ONSOD provides technical support to the bi-lateral agreement process for the establishment and operation of stations by host nations.

It should be noted that although Omega is operational, it is not completely validated. That is, there are many areas over the globe where insufficient data have been taken to verify the theoretical predictions. Also, in certain areas such as the South Pacific, with the absence of the Australian station, the accuracy which can be obtained may be substantially less than 1-2 nmi. The first phase of the validation program focused on the Western Pacific region because the U.S.-operated Loran-A stations in that area were shut down the end of December 1977. Since Omega is considered a replacement for Loran-A, sufficient data must be acquired and analyzed to demonstrate that Omega can meet the requirements now being met by Loran-A. A report on the Western Pacific validation was issued in April 1978 [39]. The next area to be validated is expected to be the North Atlantic.

A summary of the preliminary results of the Western Pacific Omega validation is given below. The assessment was based on engineering and operational data acquired throughout the area during 1976 and 1977. In addition, areas of potential improvement have been identified and the expected performance with the improvements is estimated.

The principle result of the coverage analysis was the determination of station pair combinations which provide the best navigation performance over the area. A summary of the recommended station pairs is illustrated in Figure 10. These recommendations are based on considerations of signal-to-noise ratios, geometry, and modal interference. The goal of this effort was a single coverage chart with a minimum number of regions and with the two best station pairs suitable for day and night use identified for each region. This was not possible during daytime and modal interference on the signals from Stations C and H at night. These two conditions are taken into account by the shaded regions shown in Figure 11. The shaded regions identify regions where Stations C or H can be used as an alternate during local daytime, for one of the recommended stations. In these regions, Station C or H (as indicated by the notes) provide higher signal-to-noise ratios than the respective recommended stations, but their signals undergo modal interference at night. In addition to the best recommended station pairs as shown in Figure 11, alternate stations were identified for each region to be used in the event one of the recommended stations is not available.

Another result of the alternate station analysis was an assessment of usable redundant signal coverage over the entire Western Pacific area. Redundant station coverage is marginal in some areas until the



Note 1: H can be used as an alternate for D during local daytime in shaded area.

Note 2: C can be used as an alternate for D during local daytime in shaded area.

Note 3: H can be used as an alternate for A during local daytime in shaded area.

Note 4: Alternate stations are not required in Area 4 when the recommended stations B, E, and H are available.

Note 5: C can be used as an alternate for A during local daytime.

Figure 11. Recommended LOP's for Western Pacific Area Coverage Based on Use of 10.2 kHz Only and Without Station G (Australia) Coverage.

Australian station is operational. At least one alternate station is available over 90% of the area during the night and 70% of the area during the day.

An accuracy analysis of the Omega system must consider two distinctly different, but related error sources. One is the position error resulting from phase tracking noisy signals, taking into account geometry and propagation correction (PPC) induced errors, given that the correct cycle (Omega lane) is being tracked. The other error source is caused by tracking the wrong cycle (Omega lane). The lane error results from either initializing on the wrong lane or the Omega receiver slipping into the wrong lane while tracking. The lane widths, at 10.2 kHz, based on the recommended station pairs for the Western Pacific area, vary from 8 nmi to 25 nmi. The 25 nmi occurs when using line-of-position (LOP) A-H in the vicinity east of Guam. Station C is recommended here as an alternate for A during local daytime, but not at night because of modal interference.

Given proper lane identification, the resulting position error is obtained directly from statistical analysis of the monitor data; however, the position error resulting from tracking the wrong lane is not directly obtainable from the monitor data which is in terms of corrected phase measurements. The probability of lane error is highly dependent upon Omega receiver implementation and operation.

With a manual single-frequency LOP receiver using strip chart recorders, proper lane identification is achieved by operator interpretation. Lane identification in automatic receivers is design-dependent. In addition to the difference in lane identification, the accuracies of the two types of receivers, given correct lane identification, will differ because the automatic receiver uses more than one frequency.

The accuracy, given correct lane identification, was estimated for the two types of receivers. The measure of accuracy selected for the Omega position fix is the 95% circular error. The 95% circular error is the radius of a circle including 95% of the position fixes, both day and night. The stated accuracy goal of the Omega system is 1 nmi RMS during daytime and 2 nmi RMS at night. This can be translated to a day and night 95% circular error of 4 nmi.

For the automatic Omega receiver, the estimated 95% circular error, based on the data at the sites shown in Table 5, is 6.3 nmi. For the manual receiver, using the LOP's shown in Table 5, the estimated 95% circular error is 7.3 nmi. Substantial improvements in accuracy can be achieved by propagation correction (PPC) improvements as has already been accomplished in other areas. To illustrate the potential improvement, a random sample based on 24 hourly measurements taken from the Clark BE-EH,

Table 5
Accuracy Summary
95% Circular Error

SITE	AUTOMATIC SYSTEM		MANUAL RECEIVER		
	PRESENT PPC'S (nmi)	IMPROVED PPC'S (nmi)	LOP PAIR	PRESENT PPC'S (nmi)	IMPROVED PPC'S (nmi)
KWAJALEIN	4.4	2.5	AH EH	8.4	7.5
OROTE PT.	5.4	2.1	DH EH CH EH	5.4	2.1
CLARK	7.1	3.7	BE EH	8.6	4.0
TSUSHIMA	6.7	3.0	AD DE	6.7	3.6
DARWIN	6.8*	3.7*	BE EH	8.3*	4.1*
OVERALL	6.3	3.4		7.3	4.3

*ESTIMATED

July 1976, data was analyzed. The 95% circular position error was 10.3 nmi. After PPC improvements, the expected performance should approach a 95% circular error of 3.5 nmi for the user with an automatic receiver and 4.3 nmi for the user with a manual receiver. These accuracies are compatible with the stated Omega accuracy goal.

Since a determination of the probability of identifying the correct lane could not be made directly from the data, an estimate of correct laning by the use of the 3.4 kHz difference frequency was made based on analysis of 10.2 kHz and 13.6 kHz data. The result of this analysis shows that using the present PPC's, laning errors can be expected to occur between 14% and 23% of the time. When the PPC improvements are made, the laning errors should be reduced to between 2.8% and 3.9% of the time.

Specific conclusions are:

- Coverage Assessment

  The Omega system coverage over the Western Pacific area can support enroute air and marine navigation. Care is required to select the best station combinations depending on location within the area and time of day. A recommended set of combinations is shown in Figure 11. In the event one or more of the recommended stations is off the air, alternate stations have been identified. The selection criteria will be considerably simplified by the addition of Station G in Australia. Station G should provide usable signals over the entire area at all times.

- LOP Phase Accuracy

  Analysis of the phase difference data from the monitor sites indicate LOP accuracies with hourly mean averages of about 15 CEC. Standard deviations about the means of less than 10 CEC are generally observed. The hourly mean errors can be substantially reduced by improving PPC's, resulting in a significant improvement in position accuracy and laning reliability.

- Position Fix Accuracy

  Omega position fix accuracies were determined by combining available LOP pairs. The 95% circular error ranged from 4.4 nmi to 7.1 nmi for automatic receivers and 5.4 nmi to 8.6 nmi for manual receivers using the present PPC's. PPC improvement can significantly reduce these errors since the hourly mean errors are generally larger than the variations about them.

- Loran-A Replacement

  In the areas that were serviced by Loran-A 1H1, 1H2, 2L1, 2L2, 2L3, and 2H6 chains which were discontinued on 31 December 1977, current Omega coverage from at least three stations is available. Over the Mariana Island region, redundant coverage is marginal during times when either Station E or H is off the air. Station G, Australia, will provide the needed redundancy over this region.

- Charted LOP's

  Recommended LOP pairs should be noted on Omega charts in those regions where they provide the best service. Alternate LOP's should also be noted, with an indication of the performance to be expected.

- Signal-to-Noise Ratios

  The Omega data analyzed generally supports the coverage predictions. In those cases where minor differences were found, the predictions are usually more conservative than the data indicate. The significant differences that exist, observed during NOSC temporary site measurements are:

signal strength from Stabion B is 8 to 18 dB higher than predicted at Clark, Orote Point, Port Moresby and Darwin, and signal strength from Station C is 4 to 10 dB lower than predicted at Port Moresby and Darwin.

- Modal Interference

  Modal interference is predicted from Station H, Japan, during nighttime within the sector between the 190° and 225° bearing angles from the station. Flight tests conducted, on bearings of 205° and 215° from the station, validated the predictions and showed the deepest nulls at approximately 3 and 4 Mm (1620 and 2160 nmi) from the station. Modal interference is predicted from Station C, Hawaii, during nighttime within the sector between the 190° and 295° bearing angles from the station. Amplitude data from nighttime test flights between Hawaii and Wake Island on approximately a 270° bearing from Hawaii, confirmed the existence of modal interference; however, the test data amplitude signatures from the two flights did not correlate well with each other or with the predictions indicating significant night-to-night variations in the modal structure. Navigation based on a single frequency exhibiting severe modal interference is susceptible to lane slippage; however, modal nulls at different frequencies are spatially displaced which suggests the use of two or more frequencies as a means to reduce the incidence of modally induced lane slippage.

- Multifrequency Operation

  Use of multiple frequencies gives a significant increase in the availability of position fixing over the use of 10.2 kHz only. The 13.6 kHz signal provides a better received S/N than does 10.2 kHz. Based on the data analyzed, the 3.4 kHz difference frequency does not appear to provide sufficient accuracy for reliable laning with the present PPC's. A realizable improvement in the PPC's should provide adequate laning performance.

The FAA began evaluating Omega for civil aviation use in 1965; however, Omega coverage at that time was limited and erratic because there were only three or four stations operating on an experimental basis. More recently, the FAA conducted flight tests over Alaska and the North Atlantic during 1975 and 1976 [40]. In addition, flight test evaluations have been conducted over a number of air carrier oceanic routes [41].

RTCA (Radio Technical Commission for Aeronautics) and AEEC (Airline Electronic Engineering Committee) working groups are now working toward resolving the problems associated with the introduction of Omega into civil aviation use. Minimal Operational Characteristics (MOC's) and Minimum Performance Standards (MPS) were prepared and published by the RTCA [42]. Shortly thereafter, ARINC (Aeronautical Radio, Inc.) Characteristics 599 for the Mark 2 Omega Navigation System was prepared and published by the AEEC [43]. These lead to the preparation and publication by the FAA of Advisory Circular 120-31A, "Operational and Airworthiness Approval of Airborne Omega Radio Navigation Systems as a Means of Updating Self-Contained Navigation Systems" [44], and subsequent amendment to Federal Aviation Regulations, Part 37, by adding paragraph 37.205, "Airborne Omega Receiving Equipment - TSO-C94" [45].

Omega has been in operational use on military aircraft since about 1970 and approximately 300 military aircraft are currently Omega-equipped. This number will increase as the result of the U.S. Air Force decision in 1976 to equip their long-range transport aircraft with Omega. Estimates of the number of civil aircraft currently using Omega is on the order of 100-200, but on the increase. Two American air carriers have contracts for Omega to replace their Loran-A. In addition, there are a number of foreign air lines ordering Omega.

Omega is also in use in conjunction with VLF communication signals. Two U.S. companies are supplying airborne equipment of this type of which over 1,100 sets have been sold.

Most of the U.S. Navy's larger vessels are equipped with marine Omega receivers. A significant number of marine receivers have been sold for merchant vessel high-seas navigation.

Differential Omega is a modification to the basic system which is being evaluated in order to increase the accuracy in a local area. A fixed station is used to receive the basic Omega signals, calculate the phase corrections which are then broadcast over an uplink to users within a 100 to 200 mile radius of the station (see Figure 12). Since the source of greatest error in Omega is in the phase correction predictions, differential Omega provides a significant improvement in accuracy within the area that the measured phase corrections are more accurate than the predictions. Accuracy near the differential station approaches 1/4 mile and increases with range from the station. The U.S. Navy also recently evaluated differential Omega in the Gulf of Mexico area [46]. Also, the French Navy is evaluating differential Omega for marine applications with an interest in including airborne applications also [47]. In addition, differential Omega was one of the candidate systems under consideration to meet the U.S. Coastal Confluence Zone Requirement prior to the selection of Loran-C [48-50].

VLF Communications Signals

The U.S. Navy operates a VLF communication system to provide a global, all-weather, highly redundant communications service to ships and submarines. The use of U.S. Navy VLF communications signals for navigation became possible in 1970 when the transmitting stations were phase-stabilized with atomic clocks. Although the Navy has stated that these signals are not provided for navigation purposes (and assumes no responsibility for their misuse), the Navy informed the FAA that there is no objection to the use of VLF communication for navigation provided that the stations are not assigned additional missions as navaids, and notification procedures of the U.S. Naval Observatory are satisfactory to all concerned [51]. There are 10 stations available with assigned frequencies between 16 kHz and 24 kHz. Table 6 lists their identification, location, and frequency. High power is radiated, ranging from 100 KW to 1000 KW, to assure high signal-to-noise ratios at any receiver location. Stable signals are propagated over long distances as a result of the spherical earth ionosphere waveguide phenomena similar to that experienced at the lower VLF frequencies of Omega; however, because of the higher frequencies, the incidence of modal interference is higher. Consequently, there are variations in the propagation velocity and attenuation caused by changes in the height and density of the ionosphere similar to that observed at the Omega frequencies.

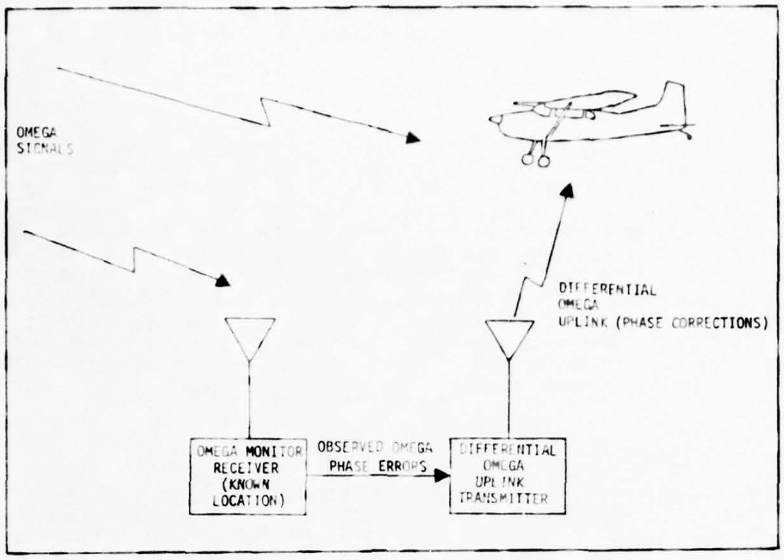


Figure 12. Differential Omega Concept

Table 6
VLF Communications Stations

IDENTIFICATION	LOCATION	FREQUENCY (kHz)
NSS*	ANNAPOLIS, MARYLAND	21.4
NAA*	CUTLER, MAINE	17.8
NBA*	BALBOA, PANAMA CANAL ZONE ...	24.0
NLK*	JIM CREEK, WASHINGTON	18.6
NPM*	LUALUALEI, HAWAII	23.4
NWC*	NORTHWEST CAPE, AUSTRALIA ...	22.3
GBR	RUGBY, ENGLAND	16.0
NDT	YOSAMI, JAPAN	17.4
JXN	HELGELAND, NORWAY	16.4
GQD	ANTHORNE, ENGLAND	19.0

*THE FORMAT OF STATIONS BEING MODIFIED BY 1978 TO TRANSMIT A MINIMUM SHIFT KEYING (MSK) AS OPPOSED TO THE PREVIOUS FREQUENCY SHIFT KEYING (FSK).

THE OTHERS ARE NOT UNDER U.S. NAVY CONTROL. ONCE MODIFIED, THE STATIONS MAY OPERATE IN ANY ONE OF THREE MODES LISTED BELOW:

(1) MSK ± 25 Hz ABOUT THE CARRIER AT 50 BAUD

(2) MSK ± 50 Hz ABOUT THE CARRIER AT 100 BAUD

(3) MODIFIED FSK, ± 25 Hz ABOUT THE CARRIER AT 50 BAUD.

THE SIMPLE DOUBLING PROCESS OF RECOVERING PHASE WILL OPERATE ON ANY OF THE THREE MODES.

Accuracy

Operational — 0.5 nm CEP/hour.

This number, quoted by one of the companies producing and distributing equipment of this type, is based on a large sample of operational observations [52]. There has not been any government-sponsored, or indepenent agency, testing of the navigational accuracy based on a thorough engineering test and evaluation approach. The reason the observed accuracy is time varying is probably because of lane slippage and inadequate phase propagation corrections.

Given that adequate phase corrections are determined and applied, the accuracy of VLF communication should be about the same as Omega. The trend in the use of VLF communiation signals for navigation is to combine the VLF communication signals with Omega signals. The average accuracy over the globe in this case should be better than Omega alone.

Coverage

A theoretical study which assessed the availability of VLF communication signals over North America and the North Atlantic concluded that signals from at least eight stations would be available based on signal-to-noise ratios [53]. That number is now reduced to seven since NBA, Balboa, has been taken off the air and placed on stand-by status. Consideration of modal interference could rule out the use of JXN, Norway, and GBR, Great Britain, over Alaska and parts of CONUS at certain times during transition.

In addition, geometry of LOP's must be considered. In regard to CONUS and Alaska, the two stations in Great Britain, GBR and GBZ, and the station in Norway, JXN, provide redundant LOP's, and in effect, most of the navigation information available from the three stations is provided by any one station.

In addition, there is a minimum radius about each of the three stations located in CONUS (NLK, NAA, NSS), within which use of that station is unreliable because of modal interference. The minimum radius, which may be different for each station, is not well established, but for planning purposes is estimated to be on the order of 300 to 600 nm.

In summary, there should be at least four VLF stations providing usable signals at any location and time over the CONUS, Alaska, and off-shore areas with two to three additional stations available most of the time to provide redundancy.

VLF Station Status

Except for scheduled maintenance, the active Navy VLF Submarine Broadcast Transmitters operate essentially continuously, to provide maximum feasible broadcast continuity. Each station conducts maintenance on a different day of the week to avoid having more than one transmitter down at any time. Naval Observatory publishes VLF maintenance schedules and frequencies to the navigation and precise time community [51]. A typical maintenance schedule is given in Ref. 52.

Reliability is defined as the percentage of time the VLF transmitters are on the air, except for unscheduled outage or casualty. The goal is 99.9 percent. The stations usually exceed 99 percent [51].

In 1975, the FAA requested the U.S. Navy to assume a navigational mission responsibility for the Navy VLF stations [54]. The request was denied [55]. Since that time, a better understanding of the use of the VLF signals for navigation has been achieved by both parties. In September 1976, the Navy informed the FAA that there is no objection to the use of VLF Communication for navigation provided that the stations are not assigned additional missions as navaids, and notification procedures of the U.S. Naval Observatory are satisfactory to all concerned [51]. In addition, the FAA sponsored an FAA/DOD/Industry meeting on 14 September 1976 to discuss the use of VLF Communication for navigation purposes [56]. A preliminary proposed Omega/VLF Approval Requirement was issued as a NOTAM and an Advisory Circular is in preparation. In May 1976, the FAA Western Region certified a VLF/Omega set as primary means of navigation (but not sole means) for ENROUTE navigation per Advisory Circular 90-45A within the 48 contiguous United States and the District of Columbia [57].

4.4 Decca

Description

The Decca Navigator is a radio position-fixing system for marine, air and land use based on continuous wave signals in the low-frequency band 70-130 kHz. Except for a special version developed for hydrographic survey work, the system is of the type in which fixed transmitting stations at known locations provide hyperbolic lines of position. The range of the system depends on various factors but is typically in the order of 240 nmi (440 km) by night and about twice that distance by day. Each user aircraft carries a special receiver which, in its simplest form, delivers the position lines as numerical "Decometer" readings which are plotted manually on a lattice chart. The intersection point of two such lines gives the position fix. Automatic and computer-based methods of reducing and displaying the Decca position fix are widely used, but many manual receivers are still in use.

A chain of Decca Navigator stations consists of a central master, and three outlying slave stations 50 to 100 miles from the master, designated Red, Green and Purple. The Red, Green and Purple descriptions refer to the respective color patterns used on Decca charts.

The three slave stations of a Decca chain are phase-locked to the common master station in the center and thus produce three intersecting patterns of position lines giving coverage in all directions around the master as illustrated in Figure 13 [58]. In practice, the user reads his lane number and fraction thereof, from the Decometers for the two patterns giving the best angle of cut at his location and rejects the third. Marine charts are generally overprinted with the two patterns appropriate to the area depicted. Figure 13 represents the plotting of a position fix.

The phase meter or Decometer cannot distinguish phase differences that are multiples of 2π but the rotor, which makes one revolution per 360° of phase, drives subsidiary pointers through gearing to indicate the number of revolutions made. The space bounded by two in-phase hyperbolae is known as a "lane" and one of the geared pointers shows a change of one lane for each revolution of the rotor. Attached to the rotor is a lane fraction pointer which sweeps a scale calibrated in hundredths of a lane.

The basic method of converting a pair of Decometer readings into a position fix (as distinct from digital data processing methods, not considered here) is a lattice or grid of hyperbolic curves superimposed upon a map or chart and numbered in Decca lane units. In general, the production of Decca lattice charts for marine navigation is the responsibility of the hydrographic authorities of the countries concerned. Lattice charts for air navigation and other uses are produced by various agencies, including the Charting Department of The Decca Navigator Company Ltd., which also prepares and supplies special charts for track plotters, flight logs and other pictorial display equipment used in ships and aircraft.

Chain frequencies are allotted according to the format shown in Table 7, which is based upon a nominal separation (at frequency 6f) of 180 Hz between the basic code values 0B, 1B, 2B, etc. [58]. Some chains deviate by 5 Hz from this separation because of factors prevailing at the time of allocation. So-called "half frequencies" 0E, 1E, 2E, etc., are spaced nominally at 90 Hz. The letters A and C denote frequencies 5 Hz below and above the B values. D and F are 5 Hz below and above the E values.

The relationship between the frequencies, lane widths and zone widths of a chain is shown in Tables 8 and 9 [58].

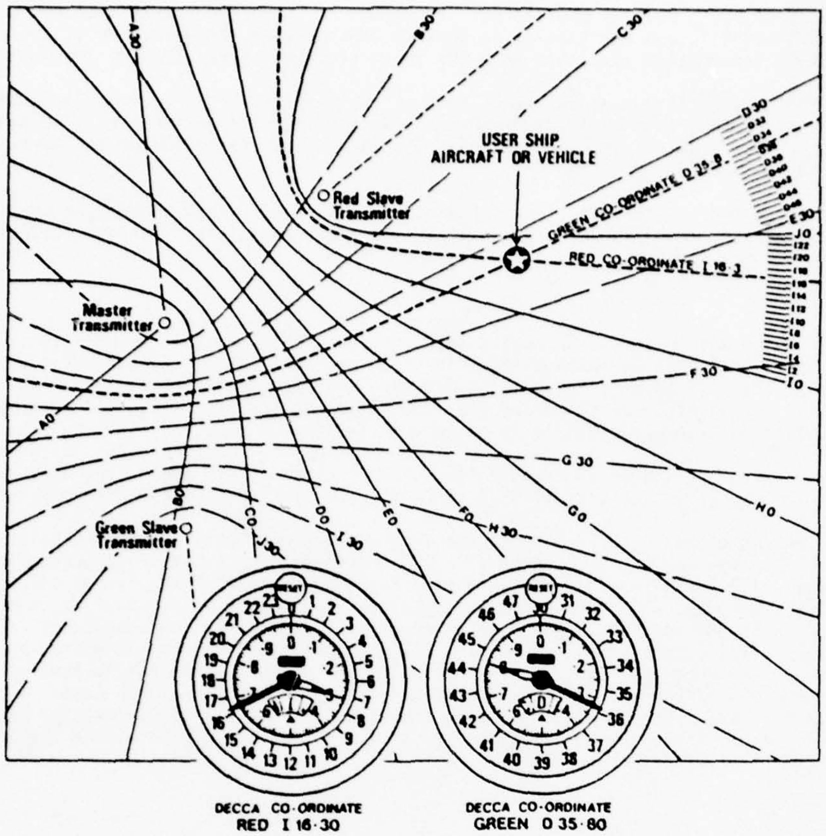


Figure 13. Decca Position Fix Plot

Table 7
Decca Chain Frequency Grouping in kHz

NOMINAL FREQUENCY (B) AND "HALF FREQUENCY" (E) GROUPS

CHAIN CODE	1f (NOT RADIATED)	5f PURPLE	6f MASTER	8f RED	8.2f ORANGE	9f GREEN
0B	14.01750	70.0875	84.1050	112.1400	114.9435	126.1575
0E	14.03250	70.1625	84.1950	112.2600	115.0665	126.2925
1B	14.04667	70.2333	84.2800	112.3733	115.1827	126.4200
1E	14.06167	70.3083	84.3700	112.4933	115.3057	126.5550
2B	14.07667	70.3833	84.4600	112.6133	115.4287	126.6900
2E	14.09167	70.4583	84.5500	112.7333	115.5517	126.8250
3B	14.10750	70.5375	84.6450	112.8600	115.6815	126.9675
3E	14.12250	70.6125	84.7350	112.9800	115.8045	127.1025
4B	14.13750	70.6875	84.8250	113.1000	115.9275	127.2375
4E	14.15250	70.7625	84.9150	113.2200	116.0505	127.3725
5B	14.16667	70.8333	85.0000	113.3333	116.1667	127.5000
5E	14.18167	70.9083	85.0900	113.4533	116.2897	127.6350
6B	14.19667	70.9833	85.1800	113.5733	116.4127	127.7700
6E	14.21167	71.0583	85.2700	113.6933	116.5357	127.9050
7B	14.22750	71.1375	85.3650	113.8200	116.6655	128.0475
7E	14.24250	71.2125	85.4550	113.9400	116.7885	128.1825
8B	14.25750	71.2875	85.5450	114.0600	116.9115	128.3175
8E	14.27250	71.3625	85.6350	114.1800	117.0345	128.4525
9B	14.28667	71.4333	85.7200	114.2930	117.1507	128.5800
9E	14.30167	71.5083	85.8100	114.4130	117.2737	128.7150
10B	14.31667	71.5833	85.9000	114.5330	117.3967	128.8500

The unmodulated transmissions occupy spot frequencies from 84.00 to 86.00 kHz for master stations and pro rata at the slave frequencies. In a manually operated receiver, chain selection simply involves turning a pair of selector controls to the required frequency code number and letter.

In the very early years of the system, lane identification was of the "V" type, in which the master transmitted a 5f signal phase-coherent with the 6f during the half-second identification period in place of the normal 5f signal from the purple slave; the receiver extracted the desired 1f frequency master

Table 8
Radiated Frequencies (Chain No. 5B)

STATION	HARMONIC	FREQUENCY (kHz)
MASTER	6f	85.0000
PURPLE SLAVE	5f	70.8333
RED SLAVE	8f	113.3333
GREEN SLAVE	9f	127.5000
ALL STATIONS	8.2f*	116.1666

*"ORANGE" FREQUENCY FOR ZONE IDENT. AND STATION CONTROL/STATUS SIGNALS.

Table 9
Comparison Frequencies and Lane/Zone Widths on Baseline
(Chain No. 5B)

PATTERN	HARMONIC	FREQUENCY (kHz)	LANE/ZONEWIDTH*
PURPLE LANES	30f	425.0000	352.1 m
RED LANES	24f	340.0000	440.1 m
GREEN LANES	18f	255.0000	586.8 m
ZONES (ALL PATTERNS)	1f	14.1666	10562.0 m

*EQUAL TO HALF-WAVELENGTH AT COMPARISON FREQUENCY, FOR THE SPECIFIED PROPAGATION SPEED (HERE 299 250 km/s)

signal as the beat note. Similarly each slave in turn sent 9f and 8f together to provide a beat note of the same frequency. The present multipulse (MP) type of lane identification has been in use since the late 1950's and derives the required 1f signal from each station by a method in which twice as much information is transmitted as in the V mode. This has the result that the MP generates a coarse pattern having greater integrity at long ranges than the fine patterns: the reverse tended to apply to the earlier method. The development of the MP technique ensured the long term viability of the Decca Navigator as a practical navigational aid.

For MP lane identification each station in turn, starting with the master, radiates all four Decca frequencies (5f, 6f, 8f, 9f) simultaneously in a phase-coherent relationship. In the receiver, the four harmonics in each such transmission are summed so as to derive a pulse train having the fundamental value f; given means of memorizing the master signal so that it can be compared with the successive slaves, this reconstituted pulse signal forms the basis of the desired f-frequency coarse pattern. The short pulse recurring at the fundamental frequency is the dominant feature of the summation waveform and has the important property that it remains stable in phase in the presence of large mutual shifts in the constituent harmonics.

The complete transmission sequence of an MP chain is shown in Figure 14. Every 20 seconds the stations transmit the MP signals in the order MRGP, together with an 8.2f component. The MP signals last 0.45 seconds and are spaced at 2.5 second intervals. In receivers which include the zone identification facility, this is based on the beat note between the 8.2f and 8.0f signals from the respective stations, giving a hyperbolic pattern of which one phase-difference cycle embraces 5 zones. Normally, the zone identification information is displayed on a separate meter on which the scale is divided into five sections, AF, GB, CH, DI, EJ. It is assumed that of the two 5-zone groups represented by these markings, the user knows the one in which he is located.

Coverage

Following the introduction of the first Decca Chain in 1946, coverage has steadily grown and by the end of 1973 there are no less than 43 chains in operation or under construction, providing an accurate navigational facility in the areas where it is most needed. Two-and-a-half million square miles are covered in North West Europe alone. In North America, four chains provide navigational coverage in the Canadian Maritimes. The waters of the Persian Gulf are completely covered by two chains providing pin-point navigation along the routes to the oil terminals for the world's tankers [58].

Further to the East, approaches to Bombay and Calcutta are serviced by two chains, and in Japan the sea areas surrounding Hokkaido and Kyushu if fully covered, while further expansion of the Decca System is planned for the remainder of Japan [58].

In the Southern Hemisphere five chains are in operation in South Africa and two in Australia.

The 43 chain locations are listed below:

23 Chains in Europe	3 Chains in Japan
2 Chains in the Persian Gulf	5 Chains in South Africa
4 Chains in North America	2 Chains in Australia
2 Chains in India	1 Chain in Indonesia
1 Chain in Bangladesh	

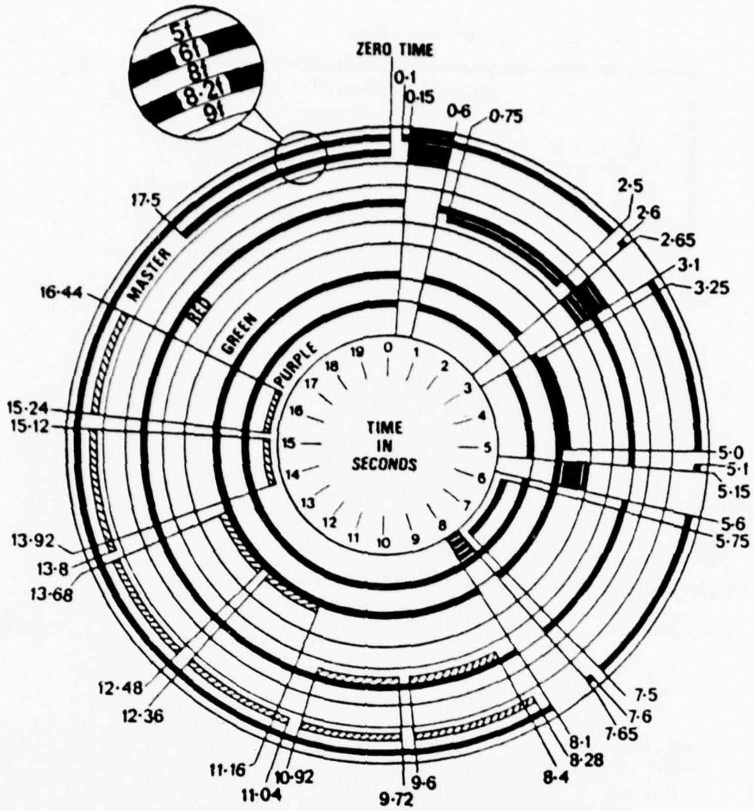


Figure 14. Transmission Sequence

The Ministry of Transport-approved coverage is 240 nmi from the Master Station. This is a minimum, and valuable use can be obtained from the system at considerably greater distances according to the conditions. World-wide coverage is shown in Figure 15.

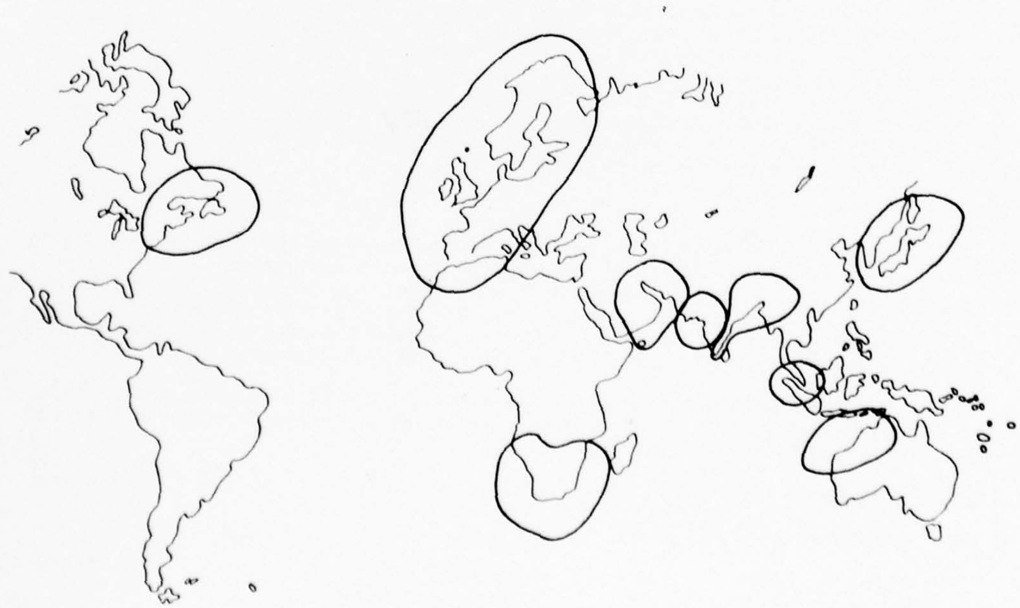


Figure 15. World-wide Coverage

Accuracy

Random errors arise from such causes as minute-to-minute changes in ionospheric conditions, short-term phase changes in the equipment, errors in readings, etc. Decometer readings taken at a given point will therefore be distributed with some degree of spread about the computed Decca coordinates of that point. A fair estimate for the standard deviation about the mean value of a large number of such readings, for daytime use of a chain at distances less than 150 miles (275 km) from the stations, and assuming seawater transmission paths, is 0.01 mean lanes. A "mean lane" is a convenient unit having a width on the baseline that is roughly the average of the red and green lane widths, namely 500 m, and corresponds to a fictitious comparison frequency of 300 Hz. A change of 0.01 mean lanes therefore represents a change of position of 5 m along a master/slave baseline [58].

The standard deviation is not wholly independent of range. During summer daylight the increase is very small, on the order of 0.003 (1 + d) where d is the range from the midpoint of the baseline in hundreds of miles. In winter this figure may be approached for a few hours in the middle of the daylight period, but is subject at other times of day to an increase by a factor of up to four or more. At night, the patterns are in general much less stable, due to increased interference by the skywave-propagated signal with the groundwave signal to which the lattice computations are related [58].

If the mean observed value of a Decca coordinate at a given point differs from the computed value for that point, and the difference remains unaltered with time, a systematic error in the pattern is said to exist at that point. Errors of this type result almost entirely from effects taking place along the paths between the transmitters and the receiver. A systematic error would result from an incorrect assumption of the mean velocity of wave propagation in computing the hyperbolic pattern. Relatively local systematic errors also occur, mainly through differences in mean propagation speed as between the transmission paths to the observer from the master and slave stations of a pair. Such uncertainties must be regarded as contributing in some measure to the random error, since the resulting pattern shifts may vary from place to place although they remain constant with time. The contribution to the random error would be greater in broken and mountainous country, for example, than over flat ground of uniform soil conductivity, and would be negligible in the case where the transmission paths lay wholly over seawater [58].

The standard deviation of 0.01 mean lanes may be taken as a reliable guide to the performance obtained in off-shore surveying with a chain sited on or near a coastline. In unfavorable terrain conditions, however (for example, in polar coastal regions where there is a juxtaposition of seawater with ground of extremely low conductivity), a standard deviation as high as 0.1 mean lanes might have to be assumed to take account of changes in fixed error from place to place. Users of the permanent Decca Navigator chains are furnished, in the data sheets, with the fullest possible details of fixed errors prevailing in the coverage of the various chains and an example of the charts used for this purpose is shown in Figure 16. While systematic errors of the kind discussed above remain, in general, unaltered with time, there is some evidence of a small seasonal fluctuation in error values in certain cases of chains sited in regions of low conductivity [58].

An example of a set of accuracy contours related to "times other than daylight" is shown in Figure 17, together with the associated table and diagram defining the time/season factor in Figure 18 [58].

4.4 VOR/DME, and TACAN

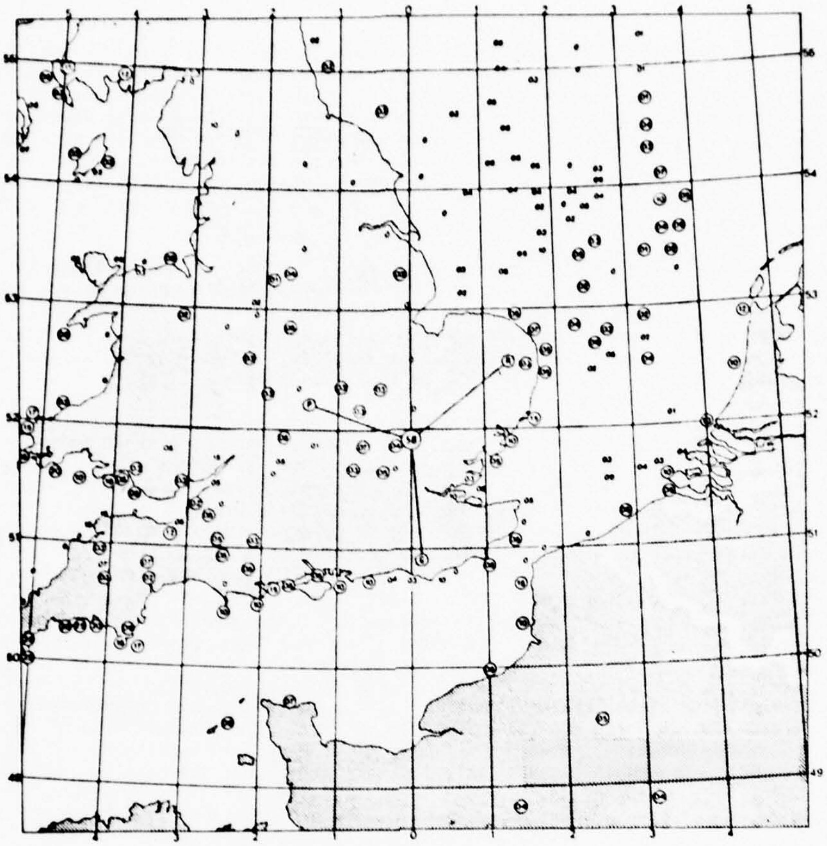
At the present time, the standard air navigation systems in the United States and many other countries are VOR (VHF omni-directional range), DME (distance measuring equipment), and TACAN (tactical air navigation system). VOR is the civil standard bearing measurement system, and DME is the civil standard distance measuring system. TACAN is a military system providing both bearing and distance information. The civil DME and the distance-measuring part of TACAN are essentially identical in all respects and are mutually compatible for civil and military users. The following paragraphs discuss these three systems and their operational characteristics in greater detail.

VOR System

The current international civil air navigation standard is the VHF omni-directional range (VOR) system. VOR is a passive, angle-measuring system which provides the user with a bearing angle relative to the VOR transmitter. In the U.S. there are approximately 1000 VOR transmitters located throughout the world. These stations are strategically located so that air routes are formed by joining radials, which are angles of constant bearing, from stations located near the desired direction of flight. Progress along the air route is determined by measuring the bearing of a crossing radial from a station that is off the airway. Significant air traffic control points along the route, called reporting points, make use of VOR intersections that are formed by radials from two VOR stations. An example of a U.S. air route is shown in Figure 19. One segment of route V187 (Victor 187) is formed by the 236° radial from the Missoula VORTAC and the 054° radial from the Lewiston VOR. (The differences between a VOR and a VORTAC will be discussed in a subsequent section.) The reporting point named Orofina is formed by the intersection of V187 and the 181° radial from Mullan Pass VORTAC which is located north of V187.

The VOR system has been the U.S. standard air navigation system since 1946 and the international standard, as established by the International Civil Aviation Organization (ICAO), since 1949. It will remain the international standard until at least 1985 and in all probability will continue to be in service until well into the 1990's.

The VOR system operates in the VHF band from 108 to 118 MHz. The band from 108 to 112 MHz is shared with the localizer portion of the instrument landing system. The spacing between adjacent VOR frequencies has recently been reduced from 100 kHz to 50 kHz. This spacing produces 200 channels, 160 of which are used for VOR and 40 for localizer systems. Since the VOR system operates at VHF, the signals are basically line-of-sight limited. This characteristic produces station range limitations at low aircraft altitudes. This can produce coverage gaps at low altitudes for locations that are not near VOR facilities. This problem can be particularly troublesome in mountainous regions where mountain



THE DECCA NAVIGATOR SYSTEM — ENGLISH CHAIN (5B)

GREEN PATTERN FIXED ERROR CORRECTIONS

CORRECTIONS TO APPLY TO OBSERVED GREEN DECOMETER READINGS TO OVERCOME FIXED ERRORS
VALUES SHOWN ARE IN HUNDREDTHS OF A LANE UNITS
FIGURES ENCIRCLED SHOULD BE SUBTRACTED
FIGURES NOT ENCIRCLED SHOULD BE ADDED

Figure 16. Example of Systematic Error Map (from "Decca Navigator Operating Instructions and Data Sheets")

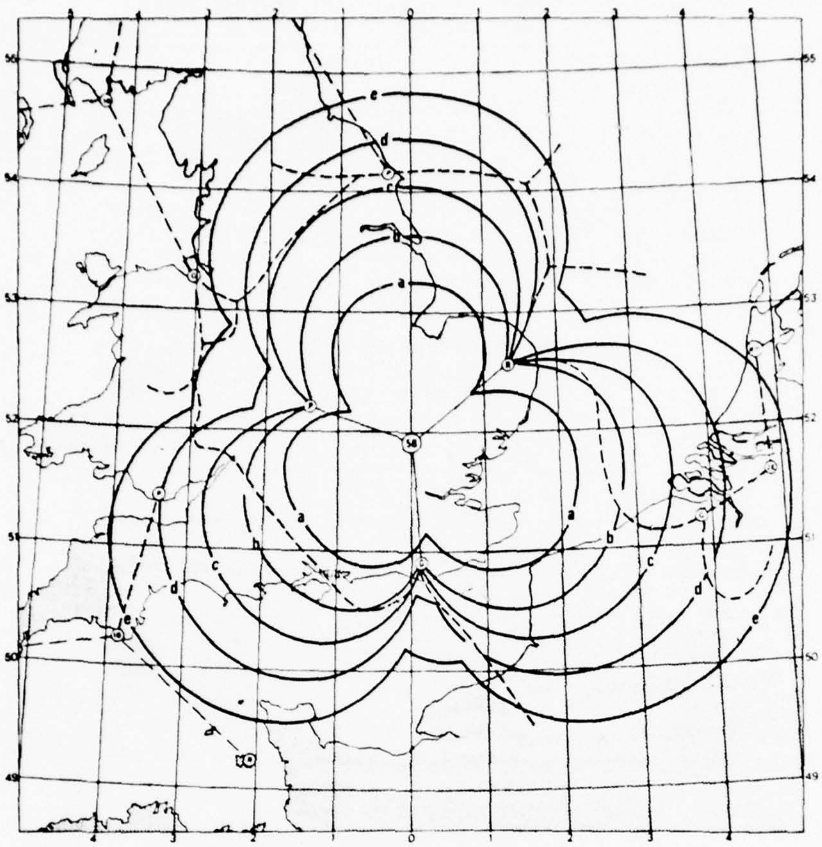
peaks can mask signals and where coverage gaps are common in the valleys. Operating at VHF can also pose serious multipath problems. The VOR facility site must be carefully selected and prepared in order to minimize multipath related problems. The results of signal multipath can produce large bends and/or scalloping (high-frequency bends) of the VOR radials which can produce errors of several degrees in the aircraft's position [34].

VOR Signal Characteristics

The VOR radiated field is composed of two parts; a figure-eight pattern which rotates at 30 Hz, and an omni-directional reference signal which contains a fixed 30-Hz tone. The phase difference between these two signals varies directly with the aircraft's bearing from the station. The two signals are in phase in the magnetic north direction. The signals are generated mechanically from the same source, an 1800 rpm constant speed motor. Thus, small speed errors affect both the rotating and reference signal and produce no apparent errors in the phase difference.

The signal is produced by a crystal-controlled transmitter. The reference phase is developed from an off-center, motor-driven tone wheel which imparts a 9960 + 480 Hz frequency shift to the 30 Hz signal. This signal is mixed with the transmitter signal to produce the frequency modulated reference phase signal. Additional modulation may be applied to this signal in the form of Morse-code tones, which identify the station and/or voice. The rotating signal is produced by feeding an unmodulated transmitter signal to a motor-driven goniometer. The output of the goniometer is the amplitude-modulated rotating phase signal.

These signals are then applied to a balanced bridge network which in turn feeds four Alford loop antennas. The Alford loop antennas are used to radiate a horizontally polarized signal. The four loops are phased to produce the figure-eight pattern for the rotating signal and the omni-directional pattern for the reference signal. The bridge networks are used to balance the antenna loads so that the two signals may be applied to the same antenna elements without affecting each other. The antennas are arranged in a square pattern and are placed a half wavelength above a metal mesh counterpoise.



THE DECCA NAVIGATOR SYSTEM — ENGLISH CHAIN (5B)

PREDICTED COVERAGE AND ACCURACY DIAGRAM (68% PROBABILITY LEVEL)
FOR TIMES OTHER THAN 'FULL DAYLIGHT'

1. See Sheet 3(a) on facing page for probable errors and time periods.
2. The time and season diagram and table are for the interpretation of the Decca accuracy contours labelled a, b, c, d or e.
3. The table gives the Variable Fixing Errors not likely to be exceeded in more than one case out of three readings.
4. Corrections to offset any known Fixed Errors are given on the following pages.

Figure 17. Example of Published Accuracy Contours for 24-Hour Operation (to be read in conjunction with Figure 16)

Airborne VOR Equipment

The airborne VOR equipment generally consists of a horizontally polarized receiving antenna connected to one of several types of receivers. Some of the receiver options include:

- panel-mounted navigation/communication (nav/com) receiver-indicator
- panel-mounted navigation receiver-indicator
- remotely mounted navigation receiver connected to panel mounted indicators

In all of these configurations the receiver function is the same; that is, to receive and detect the VOR rotating and reference signals and other audio voice and identification signals. Since the VOR and the VHF aircraft communication bands are adjacent and the VOR signal requires some audio processing, lower priced units often combine the navigation and communication functions to reduce costs; however, the high-priced units generally separate the communication and navigation functions. The AM and FM detection process produces two 30 Hz signals, one containing the reference phase signal and one containing the rotating phase signal. Comparison of the phase of these signals produces the bearing of the aircraft from magnetic north relative to the station. This phase comparison is often made by a manual phase shifter called an omni-bearing selector (OBS) which rotates through 360° of phase shift. Two indicators are used to show the 30 Hz phase relationships. The first is a course deviation indicator (CDI) which indicates the aircraft's proximity to the radial selected by the OBS control. A full scale deflection is approximately ±10° of phase shift. The second indicator is a "TO/FROM" flag. This indicator shows whether the phase between the two VOR signals is near 0° (a FROM indication) or 180° (a TO indication). This indicator arrangement permits the pilot to set an OBS bearing that is approximately equal to his desired magnetic heading (in zero wind conditions), and to fly toward or away from the VOR station. An illustration of conventional VOR indications is shown in Figure 20.

RANDOM FIXING ERRORS AT SEA LEVEL IN NAUTICAL MILES
68% PROBABILITY LEVEL

DECCA PERIOD	CONTOUR					
See Time and Season Factor Diagram below	a	b	c	d	e	
HALF LIGHT	<0·10	<0·10	<0·10	0·13	0·25	
DAWN/DUSK	<0·10	<0·10	0·13	0·25	0·50	
SUMMER NIGHT	<0·10	0·13	0·25	0·50	1·00	
WINTER NIGHT	0·10	0·18	0·37	0·75	1·50	

TIME AND SEASON FACTOR DIAGRAM

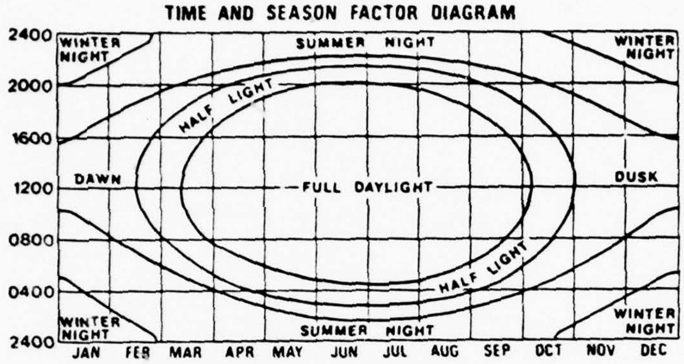


Figure 18. Table and "Onion" Time/Season Diagram for use with Figure 17. (From "The Decca Navigator Operating Instructions and Data Sheets")

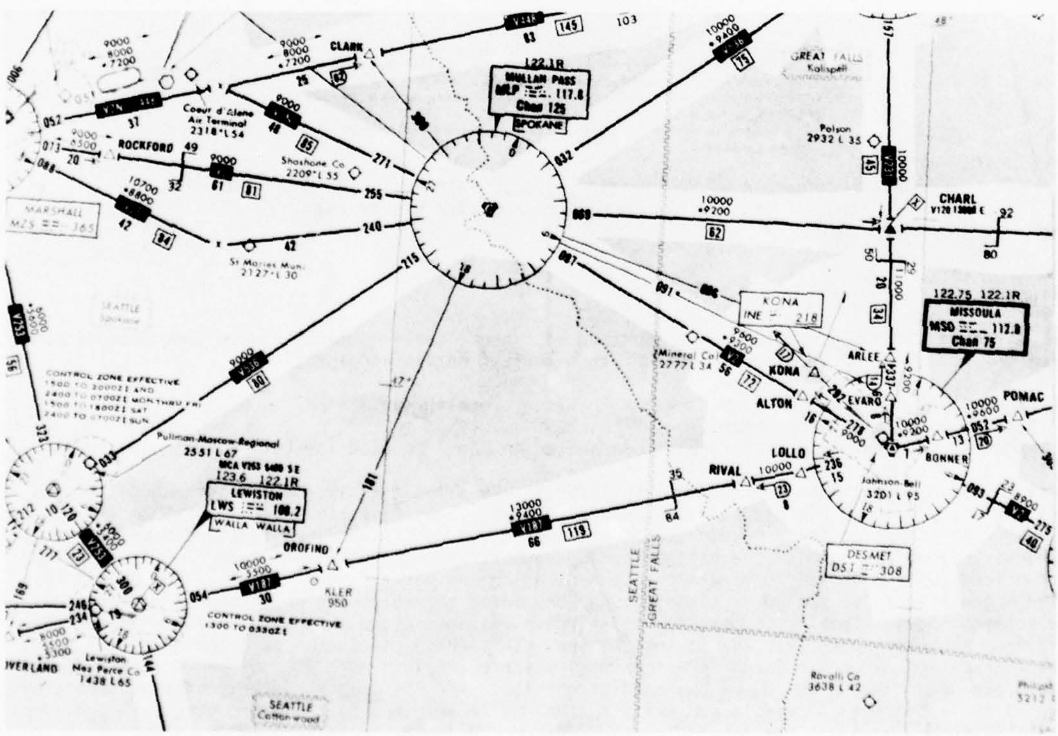


Figure 19. VOR Airways and Intersections

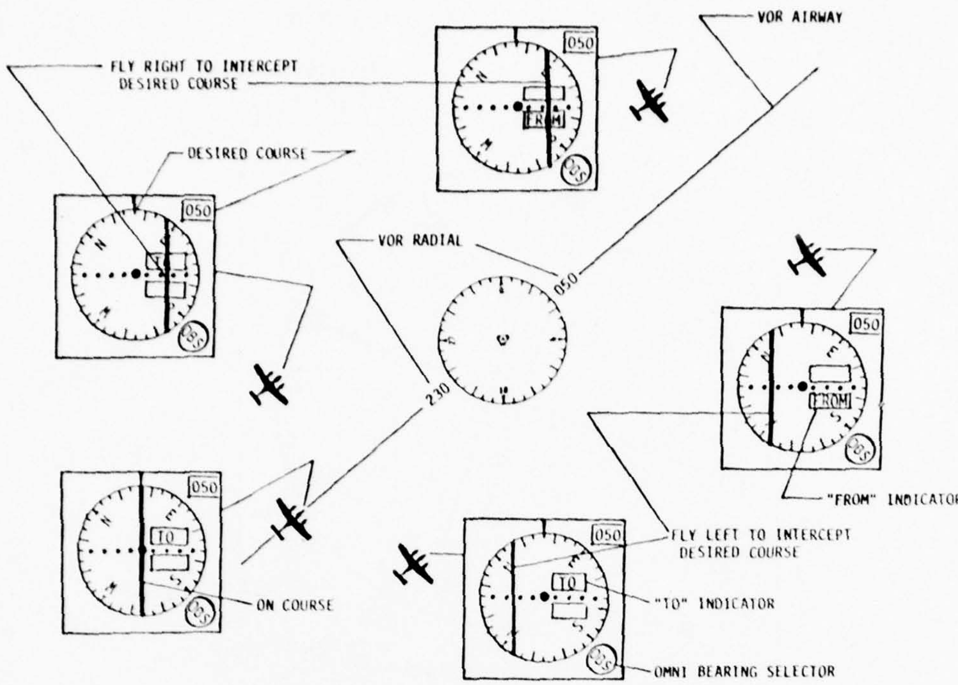


Figure 20. Course Deviation Indicator

Low cost VOR receiver-indicators generally contain only the OBS control with the CDI and TO/FROM indicator. Aircraft with an electrical source of aircraft heading information (such as a slaved magnetic compass) often combine heading and VOR information on a display called a radio-magnetic indicator (RMI). With an RMI, the aircraft heading is shown by a lubber line on a compass card. The shift between the VOR rotating and reference phase is detected by a servo-driven phase shifter and displayed, relative to north on the compass card, by a needle that pivots about the center of the compass card. In this display, the needle points in the direction of the station. The RMI display permits the pilot to quickly determine the position of his aircraft relative to the VOR station. An example of the RMI display is shown in Figure 21. Often two needles driven by two VOR receivers are provided on the RMI. One needle is connected to the VOR which provides the airway radial and the second needle points to the VOR station which provides crossing radial information so that VOR intersections or ATC reporting points may be determined readily.

In airline quality avionics, the CDI types of display and compass information are often combined in a horizontal situation indicator (HSI). In this instrument, aircraft heading, desired course and track deviation are displayed together to aid the pilot in rapidly determining the position of his aircraft relative to the desired track (OBS setting) to or from the VOR station. The course deviation signal from the VOR indicator may be used by aircraft flight control instruments for guidance purposes. Aircraft flight directors and autopilots often have the capability to use VOR course deviation inputs for aircraft flight path control.

VOR Coverage

VOR facilities are divided into three categories; they are:

CLASS	ALTITUDE	RANGE
H (high altitude)	above 45,000 ft	100 nm
	18,000-45,000 ft	130 nm
	14,500-18,000 ft	100 nm
L (low altitude)	up to 18,000 ft	40 nm
T (terminal)	up to 12,000 ft	25 nm

Limitations on VOR range are caused by signal power limitations or frequency protection. Generally, the range limitation at the higher altitude levels is for reasons of frequency protection. Additional range capability may be specified for a VOR facility if flight inspection indicates that performance is satisfactory. The existing facilities provide extensive navigation coverage and provide for air routes throughout large portions of the United States and along major air commerce routes in the world; however, large VOR coverage gaps exist at low altitudes over sparsely populated U.S. land areas, large land areas of the rest of the world and in nearly all oceanic airspace. Consequently, VOR is somewhat limited as to its area of coverage and its ability to provide navigation services in remote and over-water areas.

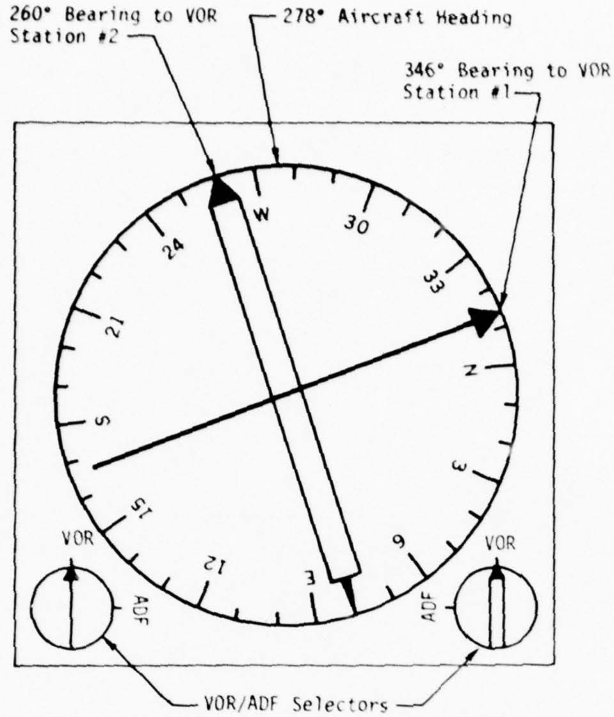


Figure 21. Radio-Magnetic Indicator

VOR System Errors

The VOR system errors are usually stated in terms of ground and airborne equipment errors, OBS setting error and flight technical error. Ground system errors are generally caused by site-related problems of multipath and reflection. Since the VOR signals are continuous wave (CW) rather than pulsed, any reflected or reradiated energy at the receiving antenna can affect the phase of the incoming signal and produce errors. These errors are minimized by careful site selection and preparation to eliminate as many sources of multipath as possible. In some areas, however, this is not possible or practical. When this happens there are three options available to the designer. They are:

- Move to another site
- Use the existing site but publish notices to airmen (NOTAM's) describing unusable VOR radials
- Install a Doppler VOR

The Doppler VOR will be discussed in a subsequent section. Typical error levels for ground stations are on the order of ± 1.9 (2σ level) [59] with many facilities able to achieve $\pm 1.0°$ (2σ level).

Errors in the airborne system can arise from a number of sources including circuit nonlinearities, temperature variations, receiver set noise, phase angle resolution, etc. Generally, lower quality units have errors on the order of $\pm 3.0°$ or less (2σ level) while well maintained, modern receivers can achieve an accuracy level of $\pm 1.0°$ or better (2σ level). In recent tests, OBS setting errors were found to be on the order of $\pm 2.0°$ (2σ level) [60]. Flight technical error, or the ability of the pilot or autopilot to keep the CDI needle centered, is often quoted as being on the order of $\pm 2.5°$. These error contributions led the route design authorities to assign a route width to VOR airways which is the greater of ± 4.0 nm or $\pm 4.5°$.

Doppler VOR

Problems with site error led to the development of the Doppler VOR system. This system is based upon the use of wide aperture antenna principles. Theoretically, the site error is inversely proportional to the antenna aperture. In conventional VOR antenna design, the aperture is approximately one-half wavelength (about 4.4 ft). The Dopper VOR utilizes 52 Alford loop antennas uniformly spaced around a 44 ft. diameter circle. These antennas are fed by a capacitive commutator which in effect simulates the rotation of a single antenna with a radius of 22 ft. Rotation is at 30 Hz with a subcarrier frequency of 9960 Kz above the carrier frequency. The carrier frequency is amplitude-modulated at 30 Hz and radiated by an Alford loop in the center of the array. The net result is an AM reference signal at the carrier frequency and a FM rotating signal at a sideband 9960 Hz above the carrier. Thus, the roles of the rotating and reference signals are reversed in the Doppler VOR but the phase relationship between the two signals is unchanged. Consequently, the operation of the airborne receiver is unaffected and Doppler VOR signals can be received and processed by using receivers designed for conventional VOR signals. In actual fact, both the upper and lower sidebands of the 9960 Hz subcarrier are transmitted. This was done to compensate for some receiver designs in order to accommodate all VOR users.

The effective aperture of the Doppler VOR array is approximately 44 ft. or 10 times the aperture of the conventional VOR antenna. Thus, a tenfold reduction in site error is theoretically possible.

Tests on previously unsatisfactory VOR sites essentially confirmed the theoretical predictions. Maximum errors with the Doppler VOR were reduced from about 3° to less than 0.5° [34,61] at several test sites.

VOR System Developments

Some experimental development work was performed on precision VOR systems (P-VOR) in the late 1960's by the FAA. A multiple-lobe structure was developed which produced a "coarse" and "fine" pattern arrangement of phase measurements for greater airborne resolution. One experiment produced a 13-lobe pattern with a theoretical accuracy improvement of 92%. With a Doppler VOR ground system and a multi-lobe signal pattern, total VOR system equipment accuracies on the order of 0.25° were anticipated. The benefits of P-VOR could not be justified by the additional airborne and ground system costs so further development plans were dropped.

One VOR system development that is being pursued at the present time in the U.S. is VORTAC modernization. It has been recently determined that it is cost-effective to upgrade most existing tube-type VOR transmitters in the U.S. to solid-state design. This program will be carried out through 1985 by the FAA.

Distance Measuring Equipment (DME)

DME [34] is an active (as opposed to passive) electronic range-measuring system. Since 1959, the VOR and DME systems have formed the ICAO standard short-range navigation system. In addition, the distance measuring portion of the military navigation system called Tactical Air Navigation System (TACAN) uses the same signals as the civil DME. Consequently, the colocation of VOR and TACAN facilities (called VORTAC) is common practice throughout the United States and many other countries.

DME Operational Characteristics

The DME frequency band runs from 960 to 1215 MHz. The band is divided into 126 channels spaced 1 MHz apart. Of these channels, 100 have been paired with corresponding VOR frequencies. At VORTAC facilities the frequencies assigned to the VOR and TACAN are selected according to this paired arrangement. Equipment manufacturers have developed frequency selectors that can tune both the VOR and the DME by selecting the VOR frequency only. This has resulted in reduced pilot workload and improved safety by reducing the number of tuning operations and minimizing the problem of setting the VOR and DME to non-colocated facilities.

Recently, to accommodate the expanded channel capacity of VOR, which resulted from reducing VOR bandwidth from 100 kHz to 50 kHz, a modified DME signal format was introduced. The existing DME format was labeled X-channel and was retained to accompany the original 100 VOR and localizer frequencies (those that are divisible by 100 kHz, i.e., 110.10, 110.20, 110.30,...,117.80, 117.90). The new DME format was labeled Y-channel and accompanies the new VOR and localizer frequencies (110.05, 110.15, 110.25, ..., 117.85, 117.95).

The DME system basically measures round-trip propagation time from the aircraft to the ground facility and back to the aircraft. The DME signal is initiated by the aircraft interrogator which emits two RF pulses spaced 12 μs apart on X-channel and 36 μs apart on Y-channel. The pulses are nominally 3.5 μs in duration. The pulses are received at the DME ground facility, delayed for 50 μs, shifted 63 MHz in frequency and retransmitted at 12 μs (X-channel) or 30 μs (Y-channel) spacings. The retransmitted pulses are received by the aircraft's DME receiver and sorted to determine valid replies. The round-trip delay time is measured, the transponder time delay is removed and the resulting delay is converted to nautical miles (approximately 12.3 μs/nm) and displayed to the pilot. Aircraft displays are generally one of three types. Low-cost systems often employ a meter-type display where an analog voltage which is proportional to DME distance is produced. Two scale factors are usually provided in the instrument. Typically these factors are about 25 and 100 miles full scale. Airline type DME indicators often employ four rotating servo-driven drums which indicate hundreds, tens, ones, and tenths of nautical miles. Recently, light-emitting diode (LED) displays have been employed in both general aviation and airline DME instrument displays. The LED display presents a digital readout that is similar to the four-drum type indicator.

Airborne Interrogator Unit

The airborne interrogator unit produces and transmits the DME pulse pair at a pulse repetition frequency varying from 5 to 150 pulse pairs per second. The higher rate is used during signal acquisition and the lower rate is used for tracking purposes after the signal has been acquired by the airborne receiver. The pulse pairs are used to distinguish the DME pulses from other potential interfering RF energy. The peak power outputs of airborne units range from about 50 to 1000 watts. The interrogator pulse pair repetition rate is purposely allowed to vary in order to distinguish that interrogator from interrogators on other aircraft.

Ground Transponder Unit

The DME ground facility receives the interrogator signal on the appropriate frequency, delays the signal for 50 μs, shifts the frequency by 63 MHz and retransmits the pulse pair back to the interrogator. The 63 MHz frequency shift is used so that use of the same circuit elements, such as antennas and oscillators, can be used for both transmission and reception. Common design practice is to utilize a receiver intermediate frequency (IF) of 63 MHz to coincide with the 63 MHz shift provided by the ground facility.

For a number of reasons it is desirable to have the ground facility transmit a fairly constant number of pulses per second. From the ground facility standpoint, the transmitter is operating at a constant average power level which is kept within its design limits. From the airborne system viewpoint, the interrogator receives a constant number of pulses to process which aids in its design. The

DME ground facility is set up to automatically adjust the receiver sensitivity to maintain a pulse pair repetition rate of about 3000 pulse pairs per second. If there are not sufficient aircraft users to provide this output rate, the ground receiver sensitivity is increased to provide noise or squitter pulses to the transmitter to maintain the desired rate. These squitter pulses are ignored by the airborne receiver.

Since the DME ground facility is active and has an upper limit of about 3000 pulse pairs per second, the DME facility can become saturated if too many aircraft attempt to use the facility. Under the assumption that most aircraft interrogators are in the track mode which operates around 20-50 pulse pairs per second, approximately one hundred aircraft may be serviced before the facility becomes saturated. Since ground receiver sensitivity is used to control the transponder pulse pair rate, the stronger signals at the ground facility will be served at the expense of the weaker signals. Thus, preference is implicitly given to aircraft close to the facility or aircraft with powerful interrogator transmitters.

The interrogator-receiver frequencies differ in X- and Y-channel operations. For X-channel operations, channels 1-63 shift the frequency downward while channels 64-126 sift the frequency upward. On Y-channel the procedure is reversed. Twenty-six (52 counting both X- and Y-channel operations) DME channels are not assigned to corresponding VOR frequencies, but are reserved as TACAN-only channels. The correspondence between TACAN channels and VOR frequencies is shown in Table 10.

Table 10
Correspondence Between TACAN Channels and VOR Frequencies

TACAN CHANNEL	VOR/LOC FREQUENCY (MHz)	INTERROGATOR FREQUENCY (MHz)	REPLY FREQUENCY (MHz)	REPLY FREQ. - INTER. FREQ. (MHz)
1X - 16X	NONE	1025 - 1040	962 - 977	-63
17X - 59X	108.00 - 112.20*	1041 - 1083	978 - 1020	-63
60X - 63X	NONE	1084 - 1087	1021 - 1024	-63
64X - 69X	NONE	1088 - 1093	1151 - 1156	+63
70X - 126X	112.30 - 117.90	1094 - 1150	1157 - 1213	+63
1Y - 16Y	NONE	1025 - 1040	1088 - 1103	+63
17Y - 59Y	108.05 - 112.25*	1041 - 1083	1104 - 1146	+63
60Y - 63Y	NONE	1084 - 1087	1147 - 1150	+63
64Y - 69Y	NONE	1088 - 1093	1025 - 1030	-63
70Y - 126Y	112.35 - 117.95	1094 - 1150	1031 - 1087	-63

*ALL VOR LOC FREQUENCIES ENDING IN .*0 CORRESPOND TO X-CHANNEL DME AND VOR LOC FREQUENCIES ENDING IN .*5 CORRESPOND TO Y-CHANNEL DME.

DME Airborne Receiver

The receiver portion of the DME interrogator must receive and decode the DME pulse pairs and convert the information into distance. The pulse string is received, filtered to remove frequencies outside the desired band and mixed with the transmitter to produce a 63 MHz signal modulated by the pulse amplitudes. The signal then is amplified using a 63 MHz IF amplifier which produces the received pulse string. The received pulses are then compared to the transmitted pulse string and a correspondence is sought between the two pulse strings. Early versions of DME processors utilized moving gates to perform the comparison and the procedure could take as long as 30-60 sec to acquire the proper range value during the initial "lock-on" phase. Recently, digital processing and correlation detection procedures have reduced the acquisition time greatly and the process now usually takes less than 5-10 seconds with some sets claiming acquisition in a fraction of a second. After acquisition is accomplished, the processor goes into track mode in which the pulse pair repetition rate is reduced from about 150 pulse pairs per second during acquisition to about 25 pulse pairs per second or less during tracking. On ILS frequencies, the pulse repetition rate is kept at a fairly high rate (75-100 pps) to achieve a higher degree of accuracy. If the DME signal is lost for a short period of time, the DME tracking circuitry goes into a coast mode during which attempts are made to reacquire the signal without initiating the acquisition mode. If the signal is not reacquired after about 10 seconds, the pulse repetition rate increases and the search is begun for the signal in acquisition mode.

DME Coverage

Single-station coverage for DME is limited to line-of-sight similar to VOR. Thus coverage gaps do occur at low altitudes and in mountainous areas. Also, interference due to frequency protection considerations occur at high altitudes. Consequently, the coverage of DME and VOR portions of VORTAC facilities is assumed to be the same (see previous section on VOR coverage). Throughout the United States most VOR facilities are actually VORTAC's. This is especially true of those facilities which comprise the U.S. National Airways System. A high percentage of these facilities are VORTAC's, and thus, the VOR and DME coverage in the U.S. is nearly identical.

DME System Accuracy

The ICAO limit for DME system accuracy is the greater of 0.5 nm or 3% of range. In actual fact, most DME systems are much more accurate than this standard. Errors of less than 0.1 nm are not uncommon with modern digital-processor receivers. The limitation on accuracy at the present time consists of pulse delay stability at the ground facility and pulse leading-edge time measurements at the aircraft receiver.

VOR/DME Area Navigation

The VOR/DME system of navigation produces airways that converge over the ground facilities. Also, these routes tend to contain bends and "doglegs" which are necessarily caused by the requirement to fly along VOR radials. In order to permit greater flexibility in airway design and shorter, straighter routes, area navigation (RNAV) equipment has been designed and built to remove some of the constraints of the VOR/DME system.

Two generic types of RNAV units are in use. The low-cost type RNAV units are called course line computers. In these systems the VORTAC station is electrically moved along a specified radial and distance to the desired position called a waypoint. Once the waypoint position has been established, the navigation procedure to fly to or from waypoints is virtually identical to flying to or from a VORTAC. These low-cost systems generally have limited waypoint storage capacity and do not have the operational and procedural flexibility of the second type of RNAV unit which is the geographic reference RNAV system. These low-cost units generally range in price from $2,000 to $15,000.

The geographic reference RNAV system permits the use of waypoints coded in latitude and longitude coordinates. These systems often have the capability to store several waypoints, even up to entire airline route structures, in mass storage devices. Distance and bearing data from the DME and VOR are digitized and processed by a digital computer into navigation information such as distance to waypoint, cross-track deviation, track-angle error, ground speed, estimated time enroute, etc. These RNAV units are often compatible with other area navigation systems and permit VOR/DME data to be processed along with air data, compass data, inertial navigation systems, Omega, etc. These navigation and flight management systems can cost from $12,000 to $150,000 depending on the storage capacity and degree of sophistication.

TACAN System

TACAN was developed as a military navigation system after the end of World War II [34,62]. Like VOR/DME, TACAN provides bearing and distance information relative to the reference facility. From a pilot's point of view, the operation of the navigation instruments and the navigation procedures are identical with VOR/DME and TACAN. The major difference in the two systems is the generation and operation of the bearing signal which is entirely different. The only differences in the distance measuring portion of TACAN and DME are caused by the inclusion of bearing information on the TACAN signal which causes some DME pulses to be suppressed, and the slight loss of signal strength on some pulses due to amplitude modulation of the pulses. Since the DME system has been discussed in a previous section, only the characteristics of the TACAN bearing signal will be described in this section.

TACAN Bearing Signal Characteristics

The TACAN bearing signals utilize the same frequency spectrum and channels as the TACAN distance signals; that is, 960-1215 MHz in 1-MHz steps. The bearing signals are passive, however, and are generated at the ground facility. Because of its higher frequency, TACAN antennas can be physically smaller than VOR antennas. This makes mobile or portable transmitters more feasible for TACAN then VOR, and thus, increases its tactical capability for military operations. The rotating and reference signal concept used for VOR is also used for TACAN; however, the means of generating and transmitting these signals differs considerably from VOR. TACAN uses a multilobe signal format to enhance bearing accuracy. This requires a coarse-fine type of measurement in the aircraft receiver. The rotating, coarse signal is generated by rotating a parasitic element around the central element of the antenna at a rate of 15 revolutions per second. This imparts a 15-Hz amplitude-modulated signal on the DME pulses that are being transmitted through the antenna. Also rotating around the antenna at the same 15-revolutions-per-second rate is a 9-parasitic element cylinder that imparts an additional 135-Hz amplitude-modulated signal. The 135-Hz signal is used for the fine bearing measurement.

The bearing reference signals are produced by sending coded pulse bursts as each of the parasitic elements is passing through north. During the north reference bursts, the DME pulses are suppressed. The reference pulse code for the 15-Hz north reference signal differs from that of the 135-Hz auxiliary reference signal. Also the codes for X-channel and Y-channel reference pulses differ considerably. The bearing reference pulse codes are shown in Table 11. The actual north reference point occurs 132 μs and 165 μs after the start of the reference bursts for X-channel and Y-channel, respectively. In addition to the north and auxiliary reference pulses, a 1350-pulse-per-second signal is produced to be used for motor speed control and station identification tones which are transmitted every 30-40 seconds.

Table 11
TACAN Bearing Reference Signals

TACAN CHANNELS	NORTH REFERENCE 15 Hz	AUXILIARY REFERENCE 135 Hz
1X - 126X	12 PULSE PAIRS 12 μ SEC BETWEEN PULSES OF PAIR 30 μ SEC BETWEEN PAIRS TOTAL TIME ≈ 350 μ SEC	6 PULSE PAIRS 12 μ SEC BETWEEN PULSES OF PAIR 24 μ SEC BETWEEN PAIRS TOTAL TIME ≈ 140 μ SEC
1Y - 126Y	13 SINGLE PULSES 30 μ SEC PULSE SPACING TOTAL TIME ≈ 368 μ SEC	13 SINGLE PULSES 15 μ SEC PULSE SPACING TOTAL TIME ≈ 188 μ SEC

Airborne TACAN Bearing Receiver

The TACAN bearing receiver usually obtains the coded bearing signals from the same antenna that is used for the DME interrogations and replies. In typical analog receivers the coded north reference bursts are decoded and sent to one side of a coarse phase comparator. The auxiliary bursts are similarly decoded and sent to a fine signal comparator. The AM 15-Hz and 135-Hz signals are detected and separated by filters. The filter outputs are sent to two servo-driven phase shifters that are mechanically linked through a 9:1 ratio gear train. The outputs of the two phase shifters drive the coarse and fine phase comparators. When the coarse comparator is out of phase by more than ± 20°, its output is connected to a servo feedback system which drives the coarse phase shifter to null the comparator output. When this has been achieved to a satisfactory level, the servo-mechanism is switched to the fine phase comparator which drives the fine phase shifter to null the fine comparator output. In recent years, digital receiver design has developed to the point where many receivers presently in use make use of solid-state components to perform the signal detection and processing.

TACAN Bearing Accuracy

The TACAN bearing accuracy can vary considerably depending upon site errors, received signal strength and receiver design. Under ideal site conditions with strong signals, the modern solid-state receivers can provide bearing accuracies on the order of 0.2° to 0.5°. Under less favorable conditions, bearing accuracies can degrade to a 1.0° to 2.0° level. In poor signal and poor site conditions, care must be taken to assure that large bearing angle errors in the 15-Hz signal do not cause a false bearing acquisition of the 135-Hz fine signal. This error can cause bearing errors of ±40°.

REFERENCES

1. "Economic Requirements Analysis of Civil Air Navigation Alternatives, Vols. I and II," Systems Control, Inc., Palo Alto, CA, April 1978, Report FAA ASP 78-3.
2. "Study of the Performance of Aids to Navigation Systems – Phase 1, Closed Loop Model of the Process of Navigation, Systems Control, Inc., Palo Alto, CA, March 1978, Report CG-D-38-78.
3. "National Plan for Navigation," U.S. Department of Transportation, November 1977.
4. ICAO Regional Plan Publications.
5. "Annex 10 to the Convention of International Civil Aviation," Vol. II, July 1972.
6. "Loran-C, Omega, and Differential Omega Applied to the Civil Air Navigation Requirement of CONUS, Alaska, and Offshore," Vols. I, II, and III, Systems Control, Inc., April 1978, FAA-RD-78-30.
7. International Convention for the Safety of Life at Sea, 1974 Message from the President of the United States Transmitting the International Convention for the Safety of Life at Sea, 1974, London, Nov. 1, 1974, U.S. Government Printing Office, 57-1180, 1976.
8. Minimum Performance Standards, Marine Loran-C Receiving Equipment Radio Technical Commission for Marine Services, Dec. 20, 1977, report of Special Committee No. 70.
9. Advanced Navigational Techniques, W.T. Blackland, editor, AGARD Conference Proceedings Number Twenty-Eight, January 1970.
10. J.M. Beuckers, "Radio Navigation in North America - the Next 25 Years," 30th Annual Meeting of the Institute of Navigation, San Diego, California, 1974.
11. "Department of Transportation Radionavigation Action Plan Survey," September 1978.
12. Report of Department of Transportation, Air Traffic Control Advisory Committee, December 1969, Vols. I and 2, Dr. L.A. Goldmuntz, President, Ben Alexander, Chairman.
13. "Radio Aids to Navigation for the U.S. Coastal Confluence Region," prepared by Polhemus Navigation Sciences, Inc., for the Department of Transportation, U.S. Coast Guard, Vols. 1, 2, 3 & 4, March 1972.
14. "Department of Transportation National Plan for Navigation, Annex," July 1974.
15. Report to Congress, "Navigation Planning - Need for a New Direction," by the Comptroller General, March 1978.
16. Dr. A. Goldsmith, "Meeting the Priority Radio Navigation Needs of the United States," Final Report, Office of Secretary, Department of Transportation, September 1973.
17. F.J. Shafer, "General Accounting Office Study of Radio Navigation Systems," presented to Subcommittee on Coast Guard and Navigation of the Committee on Merchant Marine and Fisheries, U.S. House of Representatives, March 28, 1974.
18. Hearings on H. R. 13595, FY 1975 Coast Guard Authorization, Subcommittee on Coast Guard and Navigation, March 1974.
19. DOD Flight Information Publication, "Enroute Low Altitude," April 25, 1974.
20. DOD Flight Information Publication, "Supplement Caribbean and South America," The Defense Mapping Agency Aerospace Center, Flight Information Office, Latin America, April 25, 1974.
21. "Status of Implementation ICAO Regional Air Navigation Plan Caribbean Region," Department of Transportation, Federal Aviation Administration, March 17, 1972.
22. J.A. Pierce and R.H. Woodward, "The Development of Long-Range Hyperbolic Navigation in the United States," Published in Navigation, Journal of the Institute of Navigation, pg. 51, Vol. 18, No. 1, Spring 1971.
23. "Air Navigation Radio Aids," Airman's Information Manual, Department of Transportation, Federal Aviation Administration, Part I, pgs. 1-7, February 1974.
24. "Air Navigation Radio Aids," Transport Canada, Telecommunications Vol. XXVIII, Number 1, January 1974.
25. P.C. Scotti, "Loran-C Named Radio Navigation Aid for U.S. Coastal and Inland Waters," Department of Transportation,, U.S. Coast Guard News Release No. 70-74, PA-1, May 23, 1974.
26. Jansky & Bailey, "The Loran-C System of Navigation," report prepared for the U.S. Coast Guard, February 1962.
27. S.A. Meer, "Study of the VORTAC System and Its Growth Potential," The Mitre Corporation, November 1973.
28. K.D. McDonald, "A Survey of Satellite-Based Systems for Navigation, Position Surveillance, Traffic Control, and Collision Avoidance," published in Navigation, Journal of the Institute of Navigation, pg. 301, Vol. 20, No. 4, Winter 1973-74.

29. J.M. Beukers, "A Review and Applications of VLF and LF Transmissions for Navigation and Tracking," Beukers Laboratories, Inc., presented at ION Radio Navigation Symposium, Washington, D.C., November 1973.
30. Pierce, Palmer, Watt and Woodward, "Omega - A Worldwide Navigational System," Omega Implementation Committee, P & B Pub. No. 886B, Pickard & Burns Electronics, May 1966.
31. E.R. Swanson, "Omega," published in Navigation, Journal of The Institute of Navigation, pg. 168, Vol. 18, No. 2, Summer 1971.
32. J.M. Beukers, "Accuracy Limitations of the Omega Navigation System Employed in Differential Mode," Beukers Laboratories, Inc., presented to ION Conference, October 1972.
33. J.M. Beukers, "Global Radio Navigation - A Challenge for Management and International Cooperation," Navigation, Vol. 23, No. 4, Winter 1976-77.
34. M. Kayton, W.R. Fried, Avionics Navigation Systems, John Wiley & Sons, Inc., 1969.
35. Radionavigation Journal, Wild Goose Association, 1977.
36. Loran-C System Characterization, Wild Goose Association Publication No. 1/1976, September 1976.
37. Federal Register, Vol. 42, No. 20, Monday, January 31, 1977.
38. P.M. Morris and M.Y. Cha, "Omega Propagation Corrections: Background and Computational Algorithm," Report Number: ONSOD-01-74, Department of Transportation, U.S. Coast Guard, December 1974.
39. F.G. Karkalik, G.F. Sage, and W.R. Vincent, "Western Pacific Omega Validation," Report Number ONSOD-01-78, Systems Control, Inc. (Vt), prepared for Omega Navigation System Operations Detail, Washington, D.C., April 1978.
40. J.S. Scavullo, "Probing the Airborne Omega Environment," NAFEC Technical Letter Report, NA-77-24-LR, FAA, May 1977.
41. F. Karkalik, E. Wischmeyer, "An Operational Evaluation of Omega for Civil Aviation Oceanic Navigation," Report No. FAA-RD-77-65, Systems Control, Inc., prepared for Department of Transportation, Federal Aviation Administration, Washington, D.C., February 1977.
42. "Minimum Operational Characteristics and Minimum Performance Standards for Airborne Omega Receiving Equipment," RTCA Document No. DO-164, Radio Technical Commission for Aeronautics, 19 March 1976.
43. "ARINC Characteristics 599, Mark 2, Omega Navigation System," Draft No. 5, AEEC ONAV Subcommittee, Aeronautical Radio, Inc., June 1977.
44. Advisory Circular, AC No. 120-31A, dated 4-21-77, "Operational and Airworthiness Approval of Air-Borne Omega Radio Navigation Systems as a Means of Updating Self-Contained Navigation Systems," DOT, FAA, Washington, D.C., May 1977.
45. Federal Register, Vol. 42, No. 62 - Thursday, March 31, 1977.
46. E.R. Swanson, "Differential Omega Evaluation," Technical Note 187, NOSC/TN 187, Preliminary Report, March-April 1977.
47. G. Nard, "Results of Recent Experiments with Differential Omega," Journal of the Institute of Navigation, v. 19, p. 145-158, Summer 1972.
48. E.R. Swanson, D.J. Adrian and P.H. Levine, "Differential Omega Navigation for the U.S. Coastal Confluence Region," Navigation, Vol. 21, No. 3, Fall 1974, pp. 264-271.
49. E.R. Swanson, D.J. Adrian and P.H. Levine, "Differential Omega Navigation for the U.S. Coastal Confluence Region, Part I: Overview," Naval Electronics Laboratory Center Technical Report 1905, 2 January 1974 (AD 777 041).
50. E.R. Swanson, D.J. Adrian and P.H. Levine, "Differential Omega Navigation for the U.S. Coastal Confluence Region, Part II: Requirements, Accuracies and System Definition," Naval Electronics Laboratory Center Technical Report 1905, 2 January 1974 (AD 777 043).
51. Letter from Commander, Naval Telecommunications Command, Department of the Navy to Director, Flight Standards Service, Federal Aviation Administration, 8 September 1976.
52. W.R. Tymczyszyn, "Inflight Worldwide VLF Experience Using Global Navigation GNS-500," Global Navigation Incorporated, Torrance, California, Company Publication.
53. E.R. Swanson and M.J. Dick, "Propagational Assessment of VLF Navigation Signals in North America and the North Atlantic," Report Number FA74WA1-425-2, Federal Aviation Administration, 4 February 1975, performed by Naval Electronics Laboratory Center.
54. Letter from James E. Dow, Acting Administrator, FAA to the Honorable Frank A. Shrontz, Chairman, DOD Advisory Committee on Federal Aviation, Department of Defense, April 25, 1975.
55. Letter from Frank A. Shrontz, Chairman, DOD Advisory Committee on Federal Aviation, Department of Defense, to Mr. Dow, Acting Administrator, FAA, 30 June 1975.
56. Notice 8110.24, Subject, "Omega/VLF Navigation Equipment for Domestic, Including Alaska Use, Enroute Only."
57. Supplemental Type Certificate, Number SA3192WE, Department of Transportation, Federal Aviation Administration, May 14, 1976.
58. "The Decca Navigator System Principles and Performance," The Decca Navigator Company Limited, New Malden, Surrey, England, July 1976.
59. "Minimum Operational Characteristics - Vertical Guidance Equipment Used in Airborne Volumetric Navigation Systems," Radio Technical Commission for Aeronautics, Special Committee 116-E, Washington, D.C. 20006, March 1972.
60. R.J. Adams, "Preliminary RNAV Avionics Standards (2D and 3D Area Navigation Systems)," Systems Control, Inc. (Vt), West Palm Beach, Florida, performed for the Federal Aviation Administration, Washington, D.C., Report #FAA-RD-75-178, July 1974.
61. E.L. Kramar, "The Doppler VOR - Its Foundation and Results," in Advanced Navigation Techniques, W.T. Blackband, editor, AGARD Conference Proceedings #28, Technivision Services, Slough, England, January 1970.
62. J.B. Kennedy and I. Ross, "System Analysis of TACAN and DME for Addition of Digital Data Broadcast," ITT Avionics Division, Nutley, N.J., prepared for the Federal Aviation Administration, Washington, D.C., Report #FAA-RD-73-2, April 1973.

ACKNOWLEDGEMENT

The information and data reviewed, analyzed, and compiled during the preparation of this paper were made available through the efforts of many individuals whose support is gratefully acknowledged collectively. Special mention must be made for the contribution on VOR, DME, and TACAN provided by Ed McConkey of SCI (Vt). Also, the authors wish to recognize the invaluable support provided by SCI (Vt) Technical Publications under the direction of Clare Walker.

TRANSIT -- The Current Satellite Navigation System

by

Thomas A. Stansell, Jr.
Magnavox Government and Industrial Electronics Company
Advanced Products Division
2829 Maricopa Street
Torrance, California 90503
U.S.A.

1.0 INTRODUCTION AND SUMMARY

The purpose of this chapter is to provide an in-depth review of Transit, the Navy Navigation Satellite System, from the user's point of view. After a brief system description, a spectrum of diverse applications is described, ranging from the navigation of fishing boats to guiding submarines. Next, the Transit system status and its vitality are discussed. It becomes clear that the system is exceptionally reliable and trustworthy, that the use of and the investment in Transit equipment is growing at a remarkable rate, and that the basic system is about to be improved by the additon of a new generation of NOVA satellites. From these indications and the navigation planning initiatives described in Reference 12, this author concludes that Transit will continue to provide a valuable service until at least 1995, after which phase-over to the Global Positioning System is expected to be complete.

The second half of this chapter is devoted to a technical description of the position fix process and of the factors which influence accuracy. The satellite signal structure, the meaning of the navigation message, and the interpretation of Doppler measurements are covered in detail, followed by an overview of the fix calculation process. Finally, a thorough review of the system accuracy potential and of the factors which determine accuracy performance is given.

The Transit system grew out of the confluence of a vital need with newly available technology. (See Reference 17 for a complete review.) The need was to have accurate position updates for the inertial navigation equipment aboard Polaris submarines. The new space age technology came into being because of Sputnik I, which was launched on October 4, 1957. Drs. William H. Guier and George C. Weiffenbach of the Applied Physics Laboratory of Johns Hopkins University (APL) were intrigued by the substantial Doppler frequency shift of radio signals from this first artificial earth satellite. Their interest led to development of algorithms for determining the entire satellite orbit with careful Doppler measurements from a single ground tracking station. Based on this success, Drs. Frank T. McClure and Richard B. Kershner, also of APL, suggested that the process could be inverted, i.e., a navigator's position could be determined with Doppler measurements from a satellite with an accurately known orbit.

Because of the confluence of need with available technology, development of Transit was funded in December 1958. Under the leadership of Dr. Kershner, three major tasks were addressed: development of appropriate spacecraft, modeling of the earth's gravity field to permit accurate determination of satellite orbits, and development of user equipment to deliver the navigation results. Transit became operational in January of 1964, and it was released for commercial use in July of 1967. The user population has grown rapidly since that date, as detailed in Sections 4.1 and 4.4 of this chapter, and today commercial users far outnumber government or military users. Of considerable interest is the amazing diversity of applications which will be described in Section 3.0.

2.0 TRANSIT SYSTEM DESCRIPTION

This section is a very brief description of the Transit system permitting the reader to move quickly into a review of system applications. More detailed system descriptions will be provided in later sections of this chapter.

The Applied Physics Laboratory of Johns Hopkins University (APL) has played the central role in development of Transit. The original idea was conceived there, most of the actual development was performed there, and APL continues to provide technical support in maintaining and improving the system.

At this time there are five operational Transit satellites in orbit. Figure 1 illustrates their physical configuration: four panels of solar cells charge the internal batteries, and signals are transmitted to the earth by the "lampshade" antenna, which always points downward because of the gravity gradient stabilization boom. An elongated object in orbit naturally tries to align with the earth's gravity gradient. Magnetic hysteresis rods along the solar panels damp out the tendency to sway back and forth by interaction with the earth's magnetic field; that is, mechanical energy is converted to heat through magnetic hysteresis.

As illustrated by Figure 2, the satellites are in circular, polar orbits, about 1,075 kilometers high, circling the earth every 107 minutes. This constellation of orbits forms a "birdcage" within which the earth rotates, carrying us past each orbit in turn. Whenever a satellite passes above the horizon, we have the opportunity to obtain a position fix. The average time interval between fixes with the existing 5 satellites varies from about 35 to 100 minutes depending on latitude, as shown in Figure 3. Sections 4.3 and 4.7 describe plans for additional satellites which will improve the time interval statistics.

Daily operation of Transit is by the Navy Astronautics Group headquartered at Point Mugu, California, with tracking stations located at Prospect Harbor, Maine; Rosemount, Minnesota; and Wahiawa, Hawaii. As illustrated by Figure 4, each time a Transit satellite passes within line of sight of a tracking station, it receives the 150 and 400 MHz signals transmitted by the satellite, measures the Doppler frequency shift caused by the satellite's motion, and records the Doppler frequency as a function of time. The Doppler data are then sent to the Point Mugu computing center where they are used to determine each satellite's orbit and to project each orbit many hours into the future.

The computing center forms a navigation message from the predicted orbit, which is provided to the injection stations at Point Mugu and at Rosemount. At the next opportunity, one of the injection stations transmits the navigation message to the appropriate satellite. Each satellite receives a new message about every 12 hours, although the memory capacity is 16 hours.

Unlike earth-based radiolocation systems which determine positon by nearly simultaneous measurements on signals from several fixed transmitters, Transit measurements are with respect to sequential positions of the satellite as it passes, as illustrated by Figure 5. This process requires from 10 to 16 minutes, during which time the satellite travels 4,400 to 7,000 kilometers, providing an excellent baseline.

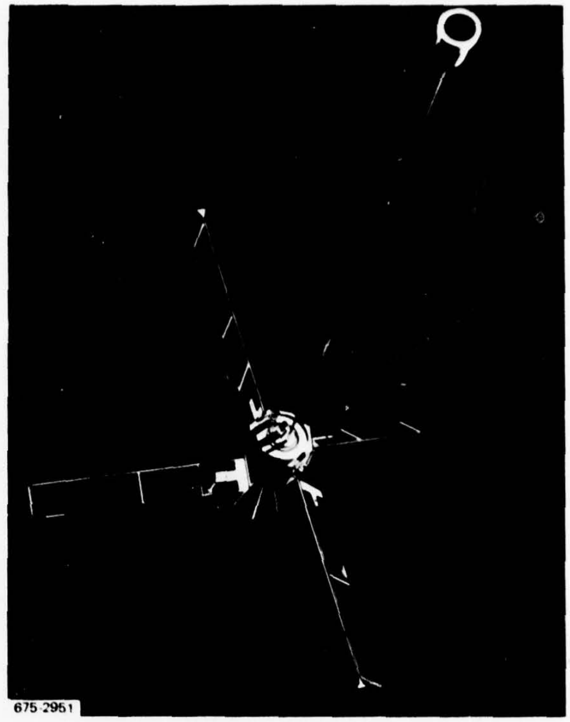


Figure 1. Physical Configuration of Transit Satellite

Because Transit measurements are not instantaneous, motion of the vessel during the satellite pass must be considered in the fix calculations. Also, because the satellites are in constant motion relative to the earth, simple charts with lines of position are impossible to generate. Instead, each satellite transmits a message which permits its position to be calculated quite accurately as a function of time. By combining the calculated satellite positions, range difference measurements between these positions (Doppler counts), and information regarding motion of the vessel, an accurate position fix can be obtained. Because the calculations are both complex and extensive, a small digital computer is required.

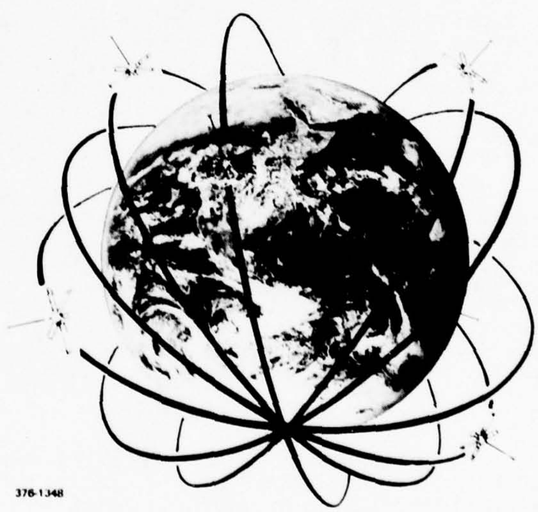


Figure 2. Transit Satellites form a "Birdcage" of Circular, Polar Orbits About 1075 km Above the Earth

Transit is the only navigation aid with total worldwide availability at this time. It is not affected by weather conditions, and position fixes have an accuracy competitive with short range radioloction systems. Each satellite is a self-contained navigation beacon which transmits two very stable frequencies (150 and 400 MHz), timing marks, and a navigation message. By receiving these signals during a single pass, the system user can calculate an accurate position fix.

There are two principal components of error in a Transit position fix. First is the inherent system error, and second is error introduced by unknown ship's motion during the satellite pass. The inherent system error can be measured by operating a Transit set at a fixed location and observing the scatter of navigation results. Figure 6 is a plot of such data from a dual-channel Transit receiver showing a radial scatter of 32 meters rms. Dual-channel results typically fall in the range of 27 to 37 meters rms. Less expensive single-channel receivers, which do not measure and remove ionospheric refraction errors, typically achieve results in the range of 80 to 100 meters rms. The second source of position fix error is introduced by unknown motion during the satellite pass. The exact error is a complex function of satellite pass geometry and direction of the velocity error, as explained in Section 6 of this chapter, but a reasonable rule is that 0.2 nautical mile (370 meters) of position error will result

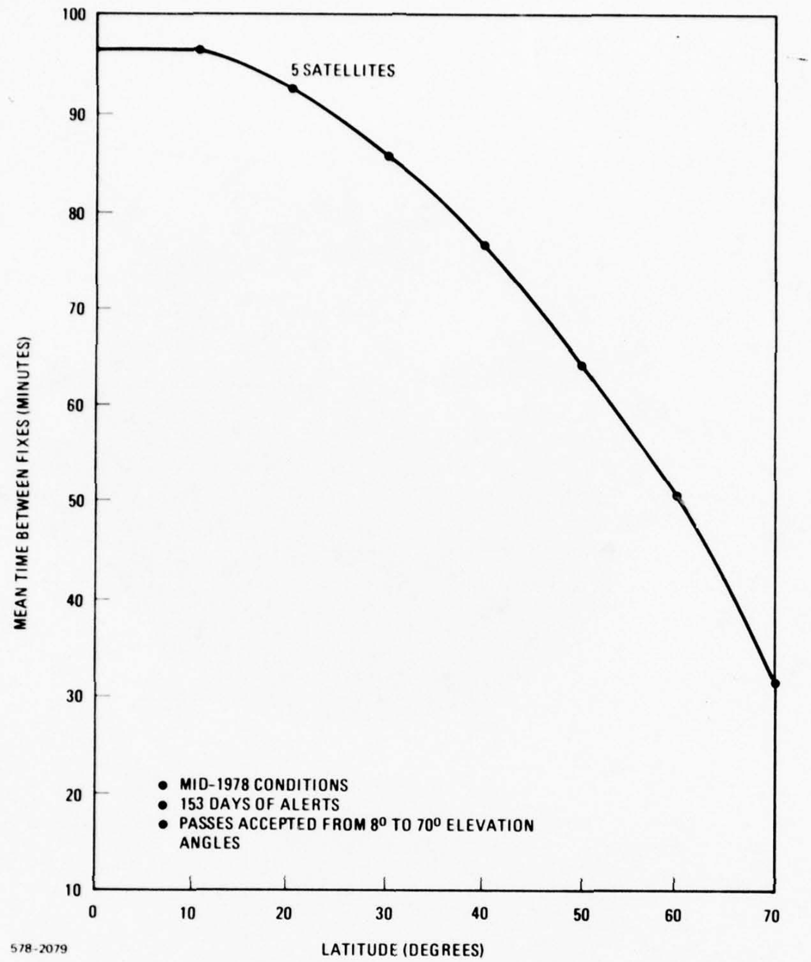


Figure 3. Mean Time Between Position Fixes as a Function of Latitude with the 5 Transit Satellites Operational in mid-1978

from each knot of unknown ship's velocity. Figure 7 is a plot of approximate position fix error as a function of unknown velocity magnitude for dual-channel and for single-channel Transit receivers. The effects of typical altitude errors and ship's pitch and roll have been included in this curve as well.

Figure 8 illustrates the preferred mode of operation for a moving navigator. Between satellite fixes the computer automatically dead reckons based on inputs of speed and heading. The dead reckoning process also is used to describe ship's motion during each satellite pass. After the position fix has been computed, latitude and longitude adjustments are applied, thus correcting for the accumulated dead reckoning error.

3.0 TRANSIT APPLICATIONS

3.1 Product Trends

The Transit system provides a combination of capabilities which cannot be obtained with any other system today. These are:

- Total global coverage
- All weather operation
- Accuracy approaching that of short range radiolocation systems
- Independence from shore-based transmitters
- Unequaled dependability

As a result, there has been a steady and dramatic increase in both the number of applications and the types of equipment available. The range of applications is truly surprising. Transit equipment is used aboard and/or for:

- Land Survey
- Fishing boats
- Private yachts

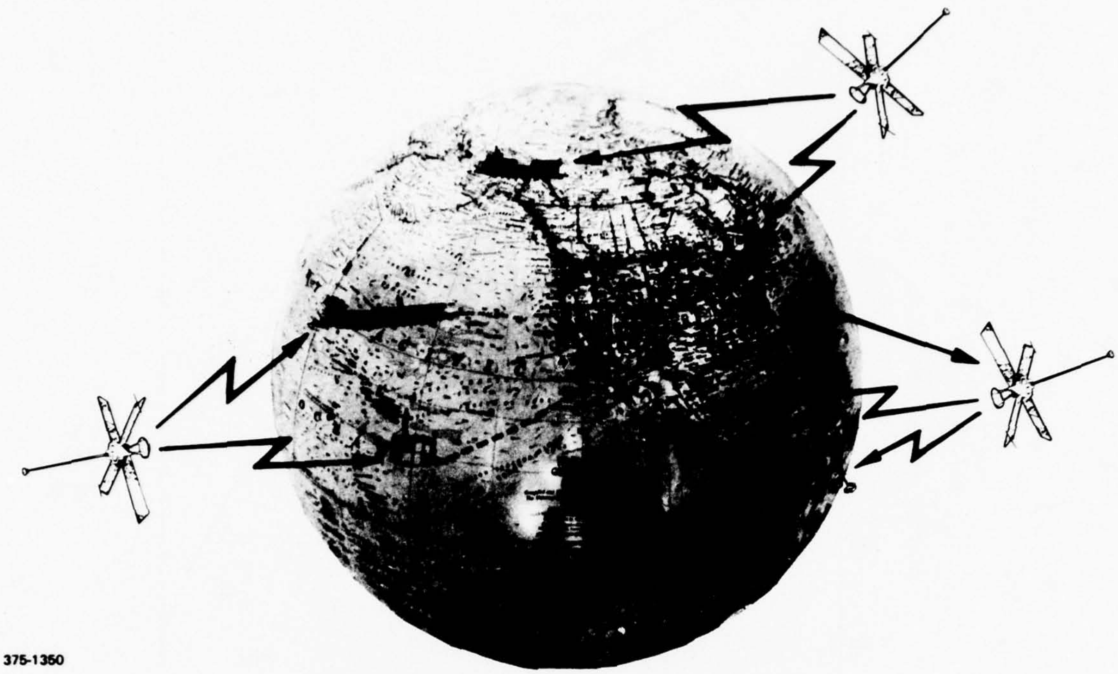


Figure 4. Schematic Overview of the Transit Navigation Satellite System

- Commercial ships (tankers, freighters, etc.)
- Military surface ships
- Submarines
- Offshore drill rigs
- Oil exploration vessels
- Oceanographic research vessels
- Hydrographic survey vessels
- Drifting buoys

To match the growing user interest and to take better advantage of available technology, Transit user equipment has evolved dramatically since the early equipment designs of 1967. Figure 9 is one view of this evolutionary process, showing the many different types of Transit receivers developed since 1967 by just one company.

Figure 10 is another view of the equipment progress, showing the evolution of Magnavox single-channel satellite navigators from 1968 through 1976. In 1968 only dual-channel receivers were available, and a minicomputer occupied most of a rack. In 1971 a single-channel receiver was introduced and by then minicomputers were only 12 inches high. In 1973 the noisy, electromechanical Teletype was replaced by a quiet and compact viedo terminal with a cassette tape reader for loading the computer program. In 1975 new technology permitted the receiver to be implemented on a pair of circuit boards which fit within the computer. Also, minicomputers became smaller, permitting greater freedom in the shipboard installation.

The final step in Figure 10 is the first production satellite navigator based on microcomputer technology, the MX 1102. In addition to being smaller, less expensive, and far more reliable than its predecessors, this new type of navigator also has more functional capability. For example, the MX 1102 not only tests itself thoroughly every two hours, but it will identify which module to replace if a failure does occur. Actual field results show a reliability of well over one year mean time between failure (MTBF). Thus, modern technology has lowered the cost and improved the capability of satellite navigation instruments.

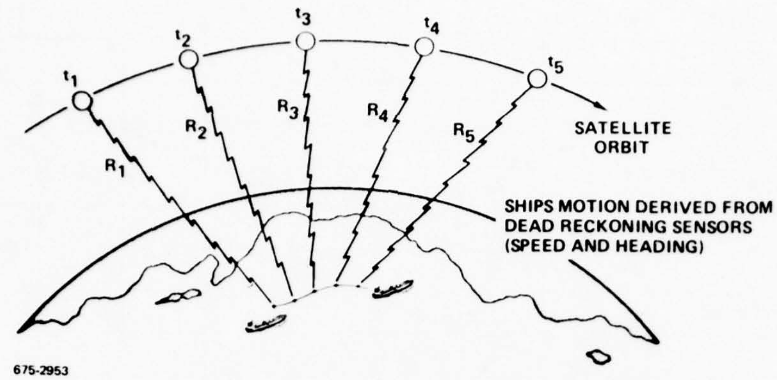


Figure 5. Geometry of a Satellite Pass

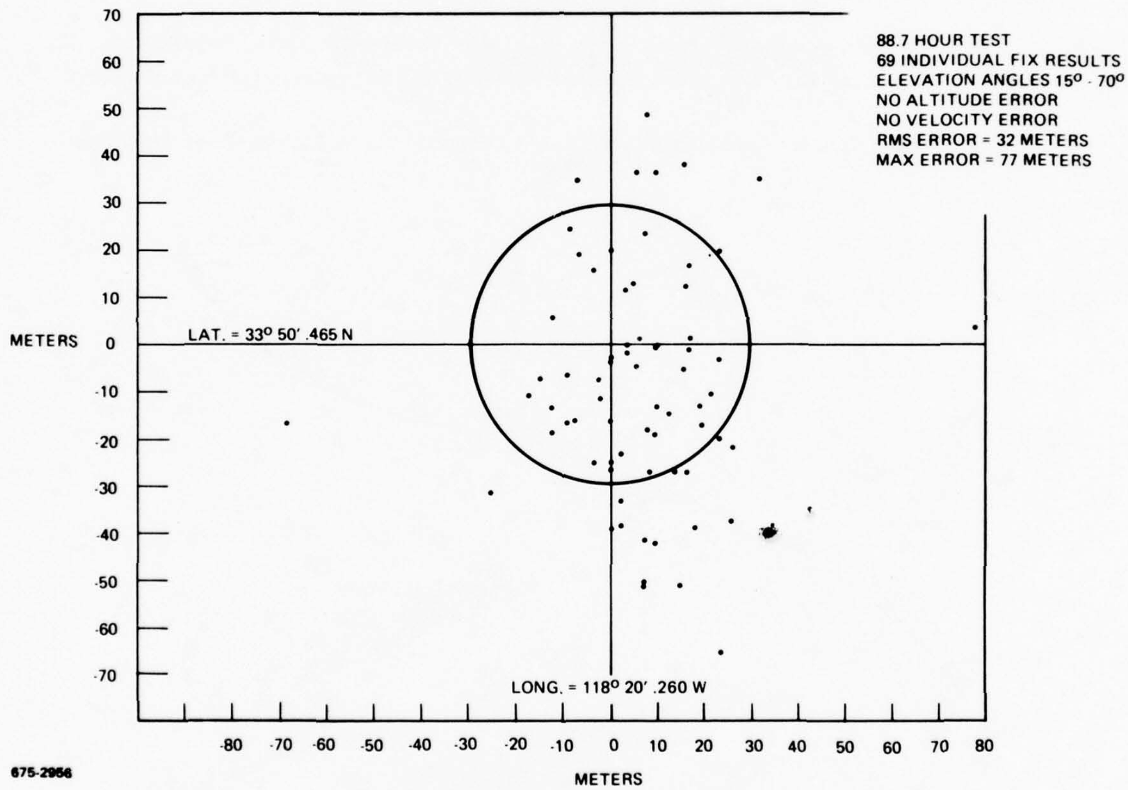


Figure 6. Typical Dual-Channel Satellite Position Fix Results

3.2 General Navigation

Because of availability of instruments like the MX 1102 shown in Figure 11, general navigation applications of Transit have dramatically increased in the last year or two. Such instruments provide a continuous display of latitude, longitude, and Greenwich mean time by continuously dead-reckoning between accurate Transit position fixes with automatic speed and heading inputs. In addition to the basic navigation functions, such systems determine and compensate for unknown set and drift, provide great circle or rhumb line range and bearing to any selected way point, determine the heading to steer to these way points, and in case of failure identify the faulty module.

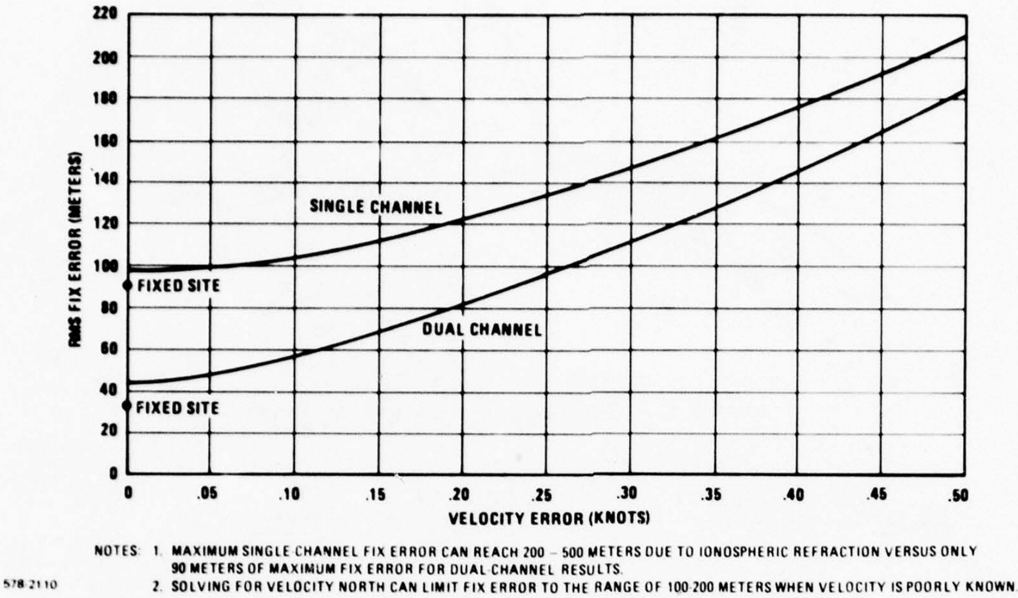


Figure 7. Approximate Satellite Position Fix Error as a Function of Unknown Velocity Magnitude

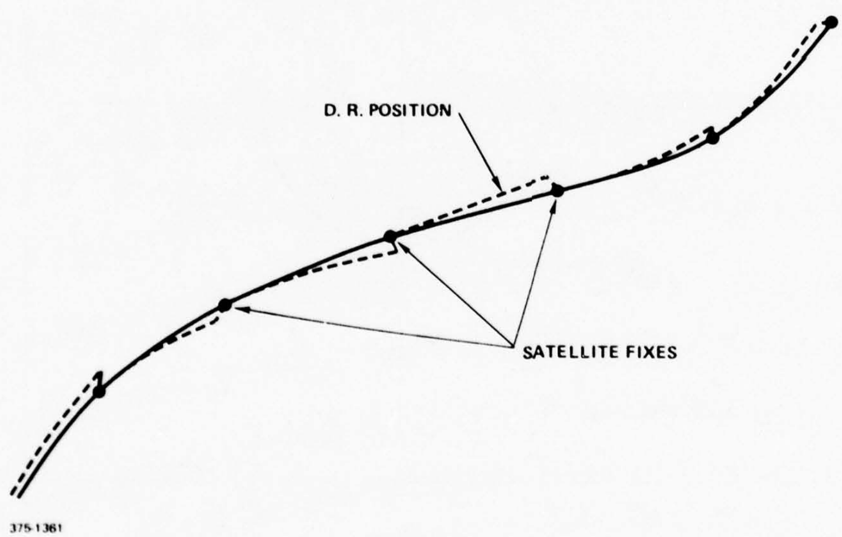


Figure 8. Dead Reckoning Error is Corrected by Each Satellite Position Fix Update

Typical applications include use aboard large fishing vessels. For example, when fishing for tuna in the southern hemisphere no other navigation aid provides the coverage or the dependable accuracy needed to assure success and to avoid fishing within 200-mile limits. Success is measured by which boat returns first with full coolers, and Transit navigation has measurably improved the rate of success.

Several large shipping companies in 1977 conducted competitive evaluations of various types of navigation equipment (Loran, Omega, and Transit, each from several manufacturers). Transit won each of these evaluations, and as a result entire commercial fleets are being equipped with Transit navigators. This trend is growing as the economic and safety advantages of dependably accurate worldwide navigation is proved over and over again. The availability of instruments with a low initial cost and with outstanding

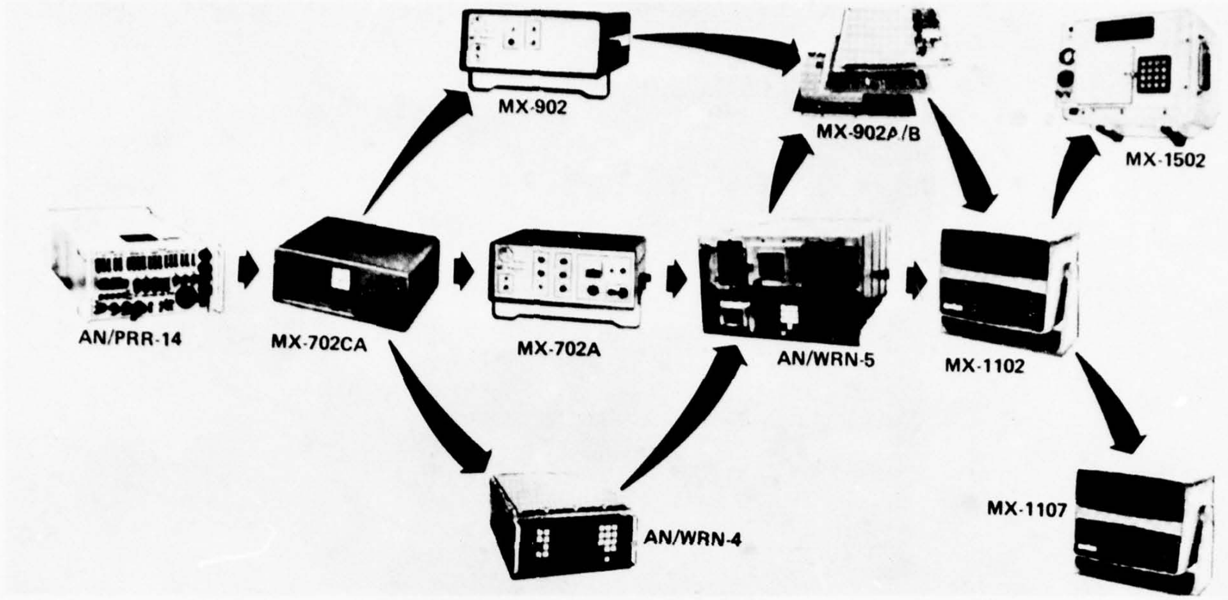


Figure 9. Evolution of Magnavox Transit Receiver Technology

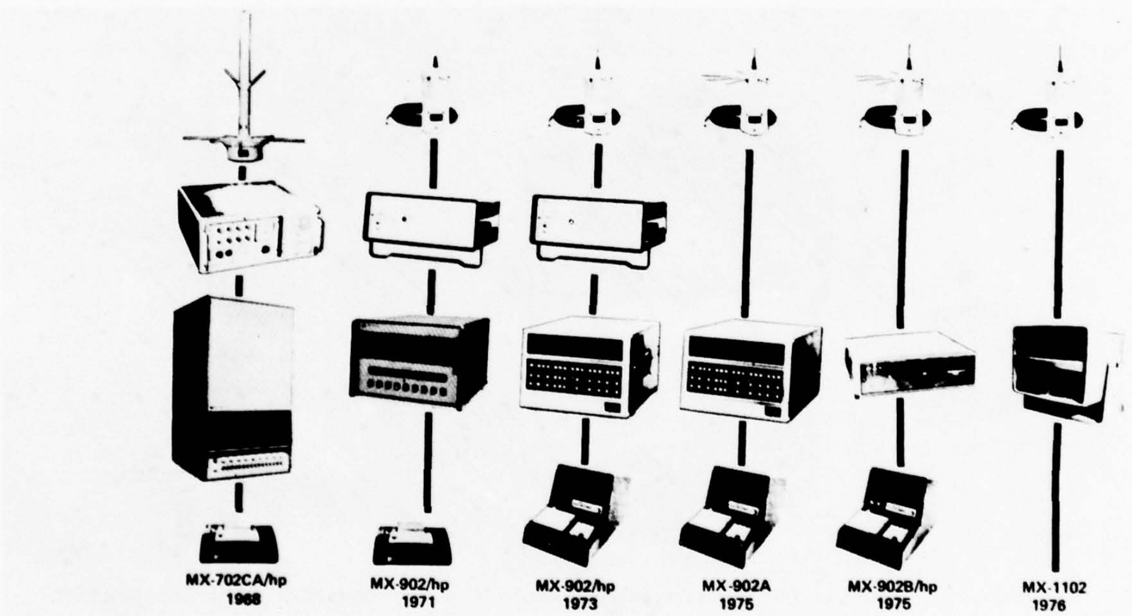


Figure 10. Evolution of Magnavox Single-Channel Satellite Navigation Equipment

reliability records, so that life cycle support costs are minimized, also has spurred the interest of major fleet operators. The need for accurate, dependable, worldwide navigation is real. For example, oil tankers passing through the Straits of Malacca truly depend on these characteristics. Often a ship will time its arrival to obtain a satellite fix just before proceeding through such hazardous waters.

3.3 Oceanographic Exploration

The first application of Transit navigation beyond its original military objectives was for oceanographic exploration. For the first time, mid-ocean scientific measurements could be tied to their geographic origin with high accuracy. The AN/WRN-4 equipment shown in Figure 9 and the equipment shown in Figure 12 are typical of the dual-channel Transit systems often used for oceanographic exploration.

In addition to the capabilities provided by commercial single-channel equipment, such as the MX 1102 of Figure 11, the dual-channel equipment gives high accuracy

Figure 11. Magnavox Satellite Navigator MX 1102

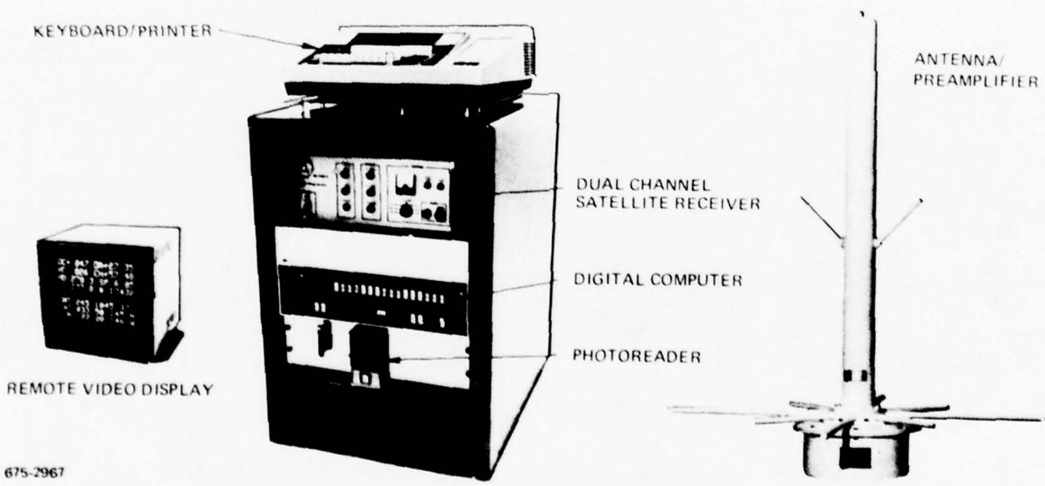


Figure 12. Typical Dual-Channel Equipment Used for Oceanographic Exploration

position fixes that are unaffected by variations in ionospheric refraction. In addition, it is typical for the system to provide a printed record of the dead-reckoned position at selected time intervals and of every satellite fix with appropriate quality indicators.

The equipment described above is now yielding to the advent of microcomputers. Figure 13 shows the Magnavox MX 1107 dual-channel satellite navigator with associated printer. This new instrument provides the same navigational accuracy capabilities as the much larger equipment shown in Figure 12.

3.4 Geophysical Survey

3.4.1 Background

In 1967 when Transit was first released for civil use, there were two immediate positive responses. One was from the oceanographic exploration community, and the other was from the offshore oil exploration community. The oceanographers were among the first civil users, but their needs have remained fairly static since the early systems were acquired. In contrast, offshore oil exploration needs have continued to grow and to become more complex.

Figure 13. Magnavox MX 1107 Dual-Channel Satellite Navigator and Printer

Prior to 1967 all offshore exploration was conducted with the aid of shore-based radiolocation systems such as Raydist, Hi-fix, etc. These systems work very well, but they have several serious problems:

- Usable range is limited, especially at night.
- The administrative and logistics costs of obtaining government approvals, transporting the equipment, installing and surveying in the shore-based stations, and operating these stations in sometimes hostile environments are very high indeed.
- Most such systems require the counting of lane crossings, so the potential for lane slips is high. This forces expensive means for occasionally verifying or correcting the lane count.

When Transit was released, there were visions of accurate, worldwide, all-weather survey operations without the time and expense required to cope with shore-based radiolocation systems. Unfortunately, simply buying a Transit navigation system did not achieve these objectives.

3.4.2 The Need for Integration

Transit provides intermittent position fixes with an individual accuracy of 27 to 37 meters, but with an additional error of about 0.2 N.Mi. per knot of unknown velocity. Survey work requires the high accuracy, but continuously. Thus, the only way to provide continuous, accurate navigation independent of shore-based stations was to combine accurate velocity sensors with the Transit fix capability in an integrated system. The first such systems were relatively crude, but very capable systems quickly evolved. Figure 14 shows a typical integrated navigation system.

3.4.3 Doppler Sonar and Gyrocompass

The first system elements to be integrated were a Doppler sonar and a gyrocompass. The Doppler sonar transmits pulses of acoustic energy to the ocean floor and evaluates the signals reflected back. The Doppler frequency shift provides an accurate measure of ship's speed with respect to the bottom in the direction of each sonar beam. Three or four beams are used to determine both fore-aft and port-starboard components of total velocity. An additional requirement is knowledge of speed of sound in water near the sonar transducer. In most cases this can be determined to satisfactory accuracy by measuring water temperature, but if salinity is likely to change drastically, a velocimeter is required for best results.

Early Doppler sonars were limited to about 200 meters of water depth before they could no longer track the bottom and had to switch to a water tracking mode, which is much less accurate. The usual Doppler Sonar today will bottom track to 300 or 400 meters, there are models available which will reach 1,000 meters or more, and systems are being developed which promise bottom tracking to maximum ocean depths.

Gyrocompasses such as the Sperry Mk-227 or the Arma Brown MK-10 compass shown in Figure 14 have been used with good success. In both cases it is important to implement automatic computer torquing of the gyrocompass to compensate for latitude, velocity, and accelerations. Not only can the computer do a better job than would be possible with the usual manual control settings, but the automatic approach avoids a major error source - the human mistake.

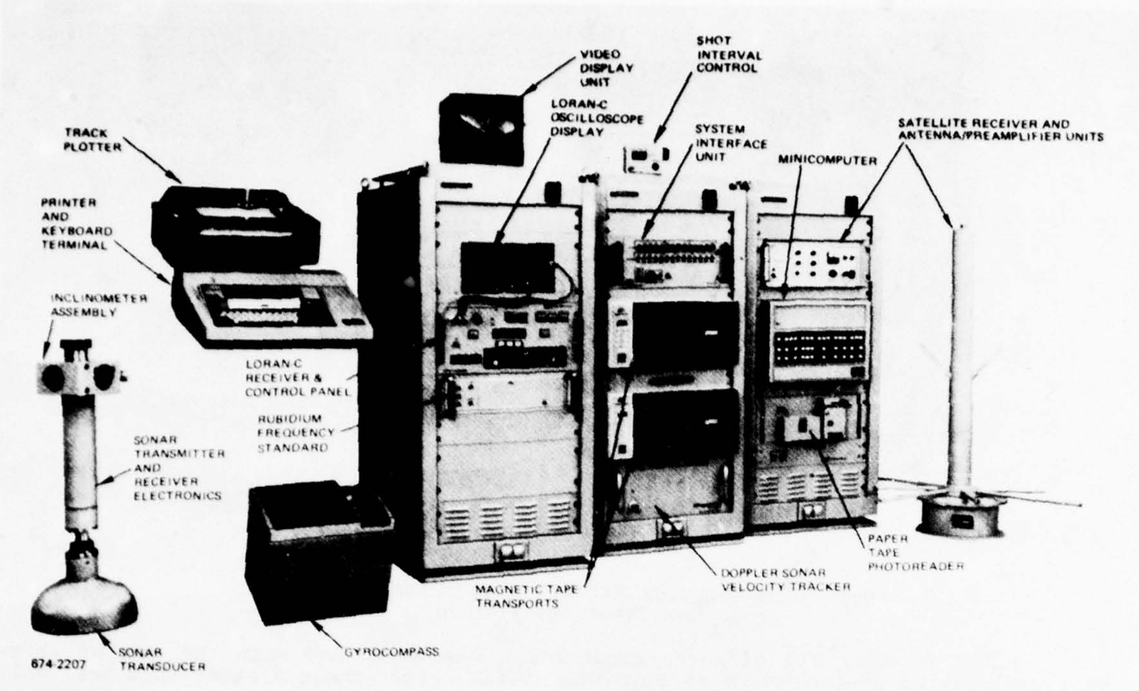


Figure 14. Typical Integrated Navigation System Components

Navigational accuracy is dependent on a number of factors, including compliment of equipment, adequacy of calibration, water depth, and sea state. Figure 15 shows how position error grows with time since the last satellite fix for a complete system and for an austere integrated system under both good and poor conditions. The figure also reveals how error increases much more rapidly when the sonar cannot reach bottom, unless a radio aid such as Loran-C is available.

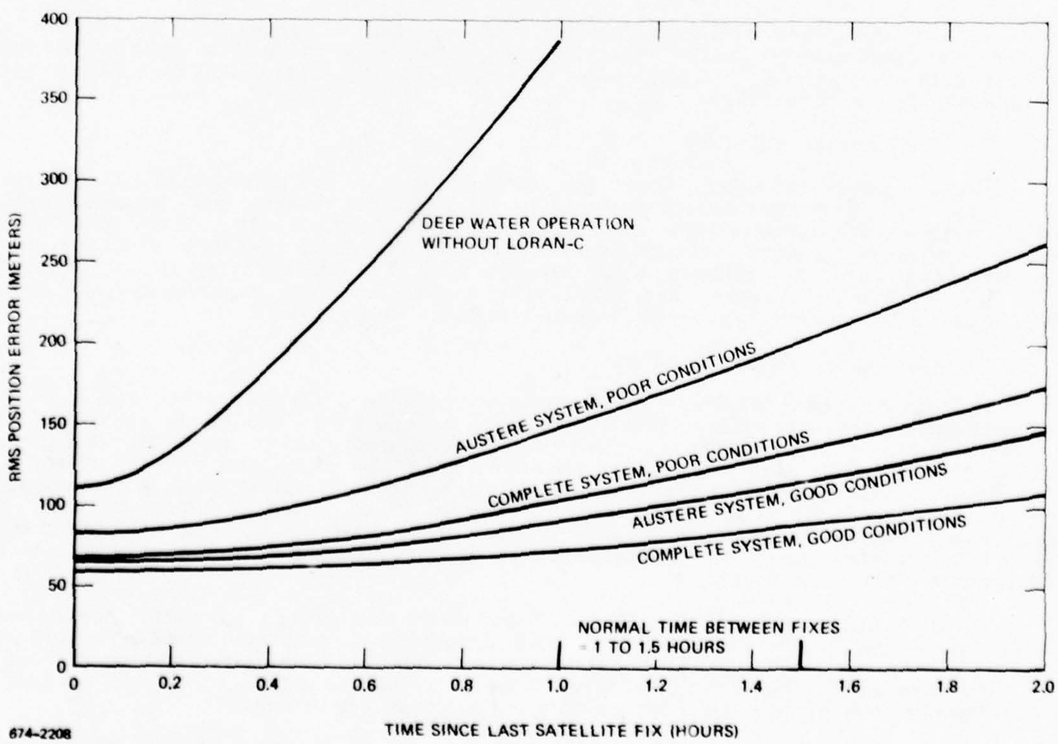


Figure 15. Integrated System Error as a Function of Time Since the Last Satellite Fix Update

3.4.4 Radio Navigation Aids

In very deep water or where it is not possible to install a Doppler sonar, some other source of velocity is needed. In many cases various radionavigation signals may be available already. For example, Figure 14 shows Loran-C components as part of the integrated system. Loran-C alone would not have sufficient accuracy because of secondary phase errors and often because of poor "crossing angles". However, by integrating Loran-C with other system elements, excellent accuracy can be obtained. Satellite fixes provide a precise geographic position reference and provide calibration of local Loran-C secondary phase errors. By having a gyrocompass and a Doppler sonar, even if in the water track mode, ship's maneuvers can be determined accurately. This permits the Loran-C readings to be filtered heavily, thus removing most of the random noise. In effect Loran-C is used to correct for the effects of unknown set and drift. Furthermore, because satellite fixes are available to provide an accurate position reference, it is possible to use Loran-C in the delta-range measurement mode with respect to a rubidium or cesium clock. Because delta-range measurements can be made on each Loran-C signal independently, useful information can be obtained with only one or two Loran-C signals, which greatly expands the area of accurate coverage and reduces the problem of poor crossing angles.

The same concepts can be used with a wide variety of radionavigation systems. When integrating with short range, high accuracy systems, a speed sensor is not needed, and the satellite fix capability is used to verify and resolve lane counts. Systems have been implemented with such radionavigation aids as: Decca Navigator, Hi Fix, Raydist, Toran, Argo, Miniranger, Trisponder, and others. Each one has its advantages, so flexible hardware and software is provided for rapid configuration with any appropriate radio navigation sensor.

3.4.5 Acoustic Transponders

One of the most sophisticated versions of an integrated system employs acoustic transponders (see References 14 and 15). The ship is equipped with an interrogater/receiver set. Every few seconds the interrogator sends out an acoustic pulse at a specific frequency. Transponders which have been placed on the bottom and are within range receive the interrogate pulse and respond by sending a pulse of their own at an individual frequency. The receiver on the ship picks up and identifies these replies and measures the total round trip delay. Such measurements, scaled with an appropriate estimate of speed of sound in water, define the range to each transponder. If the position of each transponder is known accurately, then a navigational accuracy of 2 to 10 meters can be achieved typically over an area of 3 to 10 square kilometers with only a few bottom transponders. Such systems are being used for site surveys and for precise drill rig positioning during the final approach. Although expensive, it may be the only way to achieve the required accuracy for 3-dimensional seismic surveys as well.

In the previous paragraph there was a big "if"; if the position of each transponder is known accurately. This is the difficult part. Special software has been developed to determine the transponder positions with great accuracy and in a minimum of time. The first step is to collect transponder range readings while following a specific pattern around each transponder location. Because the equations must be solved iteratively, these data are recorded in memory and used over and over until the total solution converges and the relative position of each transponder is known accurately. This technique saves time by requiring the ship to traverse the area only once; the computer does all the work after that.

Once the relative transponder positions are known, it is often necessary to determine their true latitude and longitude positions as well. This is achieved with the aid of multiple satellite position fixes. Motion of the ship relative to the transponder net can be determined accurately, but the position (translation) and azimuth (rotation) of the net are unknown. Again, an iterative solution is used in which each satellite fix improves knowledge of the net position and azimuth. As knowledge of net azimuth improves, the measure of ship's motion becomes more accurate. Such iterations are best done with all raw satellite and transponder data recorded on magnetic tape, and the technique has proved to be extremely effective and accurate.

3.4.6 Integrated Navigation System Functions

A wide variety of integrated navigation systems have been developed and deployed to aid offshore exploration. However, navigation is just one of the three major functions of an integrated system. The other two are survey control and data logging.

The system helps control the survey, for example, by firing seismic shots at defined increments of time or of distance traveled. In some installations the system actually controls steering of the vessel along the desired survey path.

Data logging is the third necessary ingredient. Unless the position at which the geophysical data were acquired is recorded, the data are worthless. Therefore, data logging must be extremely reliable with adequate tests to verify that the desired data are being recorded properly.

3.5 Fixed Point Positioning

3.5.1 Applications

All of the equipment described to this point is intended for navigation of moving vessels. Satellite signals also are being used very effectively for fixed site surveys. For instance, after a drill rig has been securely anchored, its final position must be established with the best possible accuracy.

As you would expect, surveys are conducted with satellite equipment only when this approach is cost effective. As the price of equipment has dropped and as knowledge of its advantages have spread, the number of applications has increased dramatically over the past few years.

An early application was to establish bench marks for remote sensing (aerial) surveys. For example, aerial mapping of the Amazon basin with side-looking radar has been tied to geographical coordinates by means of corner reflectors on towers placed every few kilometers. Satellite survey equipment was used to establish the position of each tower.

In general, satellite survey is used in places that are difficult to reach, where conventional survey techniques are too slow and too costly, or where local control has not been well established. For large areas without adequate control, survey by satellite is ideal. National governments such as Australia and Canada are actively involved in establishing control markers by means of satellite survey equipment.

3.5.2 Computational Techniques

Two techniques for accomplishing fixed point positioning by satellite have been developed. The first is called "point positioning". In this mode a Transit satellite receiver tracks and records every available pass at a given location. These data are fed to a computer program which combines all the raw data to obtain the single best position fix result in 3-dimensions (latitude, longitude, and altitude). The geodetic reference for such a position determination is provided by the satellite system itself.

If a reference station can be occupied within several hundred kilometers of a survey site, a technique called translocation can produce greater accuracy in less time. To implement the translocation technique, two or more satellite receivers are used, one at the reference site and the other(s) at the survey site(s). By tracking the same satellite passes, improved accuracy is achieved because the computer solves for differential position between the two points, which is not affected by common error sources.

The U.S. Government conducts many surveys with Transit satellites. The instrument normally used is the AN/PRR-14 Geoceiver shown in Figure 16. For example, adjustment of the North American Datum which is now underway depends heavily on results obtained with the Geoceiver at many survey points across all of North America. In reducing Geoceiver data, the Government has an advantage not available to the private user. This is postcomputation of each satellite orbit based on data from tracking stations taken concurrently with the survey. The result is a "precise ephemeris" orbit definition.

3.5.3 Equipment

Several different types of portable survey equipment have been developed. The original, which is still in wide use, is the AN/PRR-14 Geoceiver shown in Figure 16. On the left is the four-frequency receiver (which tracks both Transit and GEOS satellites), at the center is the antenna and preamplifier on a tripod, and on the right is the paper tape punch, which was the most reliable data recording device when the design was completed in 1967. Magnavox has delivered 55 Geoceivers which are used primarily by the U.S. Defense Mapping Agency for geodetic survey work. The Geoceiver has earned an enviable reputation for accuracy and for reliablilty.

Figure 17 shows the latest Magnavox instrument intended for fixed point survey. It is called the MX 1502 Satellite Surveyor. Being compact and lightweight, it can be transported easily. In the field it will operate for about three days on a 12-volt automobile battery. During this time, the raw data from all satellite passes will be recorded on a magnetic tape cassette. The cassette can be processed by a computing center for either point positioning or translocation results.

The MX 1502 does far more than than simply record satellite data. It computes and displays a 3-dimensional position fix result while in the field. This result often may be adequate without post-processing the tape cassette, but in any case it is extremely valuable in verifying proper system operation and assuring that the desired location has been occupied. In addition, the computed results help the surveyor to know when sufficient data have been gathered so that he can move to the next site with assurance. Assurance is a key ingredient of any survey system. Too often data are reduced to find that something was wrong and that the site must be reoccupied at great expense. The MX 1502 includes a thorough self-test capability to assure proper operation. If the self-test function detects a problem, the specific module causing the problem is indicated. Repair by replacement of plug-in modules allows the survey to continue with minimum disruption. Furthermore, after each record is placed on magnetic tape, it is immediately read back to assure no recording mistake. If an error is detected, that portion of data is re-recorded, always assuring that the proper data are recorded correctly.

The MX 1502 can learn the orbits of all Transit satellites by reading a previously recorded tape cassette. Thereafter, it will automatically go into a minimum power mode between satellite passes to reduce battery consumption, waking up just in time to track only the desirable passes. This type of new equipment will further expand the application of satellites, both for marine and for land surveys.

3.5.4 Point Positioning Accuracy

A single satellite pass can be used to obtain a latitude and longitude position fix result. As described in Section 6 of this chapter, altitude must be defined, and an error in altitude can affect the position fix accuracy. However, by processing multiple satellite passes at one location, a 3-dimensional (latitude, longitude, and altitude) position fix can be determined. The best way to do this is with a computer program which determines the one 3-dimensional position that best fits all of the Doppler measurements obtained from all of the satellite passes taken at that location. Figure 18 shows how a typical 3-dimensional survey converges toward the final answer. Each time another satellite is tracked, its data are combined with all previous data and a new solution is computed. The figure indicates that with each successive satellite pass, the latitude, longitude, and altitude parameters converge toward the final solution.

By conducting multiple convergence tests at the same location, one can determine the repeatability of the final solution. As expected, repeatability improves as the number of passes processed in each solution increases. Figure 19 makes this clear. The number beside each dot indicates how many passes were used for that position fix, and it is evident that, for example, there is more scatter to the 10-pass solutions than to the 25-pass solutions. As tabulated on the figure, the horizontal positioning repeatability is about 7 meters rms with 10 passes and about 5 meters rms with 25 passes.

Figure 19 also illustrates another important concept, which is that the position fix result is dependent on how the satellite orbits were determined. Most of the data shown in the figure was obtained before December 1975. In that month, the U.S. Navy changed the basis for computing satellite orbits from one model of the earth's gravity field to another (from APL-4.5 to WGS-72). The two points which are circled in the upper right of the figure were determined with data taken after the conversion. Thus, we can use the term "accuracy" only if we accept the satellite system as the basic geodetic reference. Otherwise, it is proper only to describe the repeatability of such a process.

The results just described are available to every system user with the necessary equipment and computer program. The principal source of error is misknowledge of satellite orbital position, made worse by the fact that orbit parameters in the satellite memory are a prediction of its position based on past tracking data. The prediction is obtained by numerical integration of the equations of motion, taking into account all known forces acting on the satellite, such as the gravity fields of the

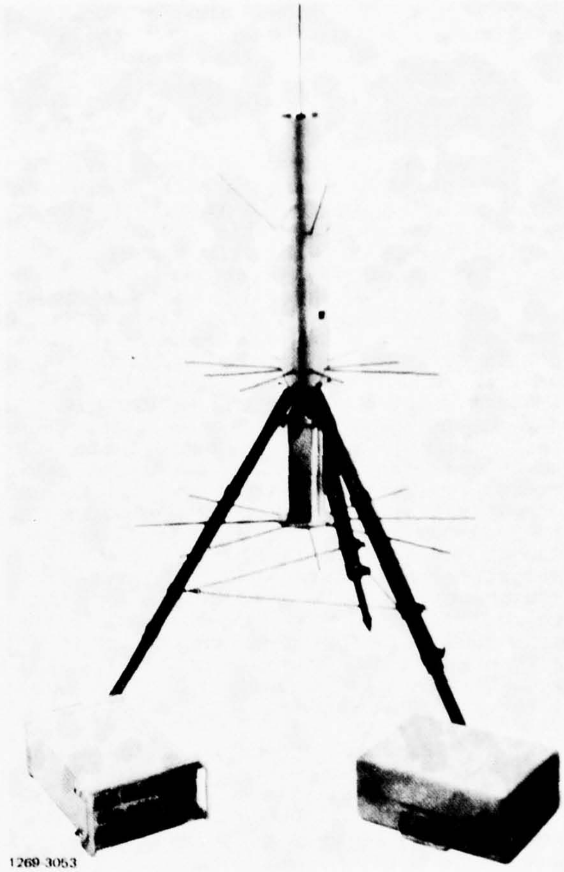


Figure 16. Original AN/PRR-14 Geoceiver

Figure 17. Magnavox MX 1502 Satellite Surveyor

earth, sun, and moon, plus drag and radiation pressures. To the extent that these forces are not known precisely, the predicted orbit will deviate from the actual orbit. These differences account for most of the 27 to 37 meters rms of error in individual Transit position fixes.

If the orbit did not have to be predicted into the future, a more precise determination could be made, and the U.S. Defense Mapping Agency (DMA) employs this technique in reducing satellite Doppler data from survey receivers such as the AN/PRR-14 Geoceiver. Field data are recorded on tape and returned to a computing center for evaluation. There the Doppler data are combined with a precise ephemeris of satellite positions based on tracked rather than predicted orbits; thus individual position fixes have a typical scatter of only 6.3 meters rms. Naturally a 3-dimensional, multi-pass solution converges to the required resolution much faster with this technique than when using predicted orbit parameters from the satellite. However, the DMA seldom computes a precise ephemeris for more than one or two satellites at a time, and immediate results cannot be obtained in the field, offsetting slightly the advantage just described. Even so, equipment using the predicted orbits must remain on station from 4 to 10 times longer than equipment using the precise ephemeris for equivalent accuracy results. The DMA has shown 3-dimensional results with 1.5 meters per axis repeatability after 25 precise ephemeris passes.

Precise ephemeris information is not available for commercial use. However, there is precedent for the DMA to supply this information to other nations based on cooperative international survey agreements.

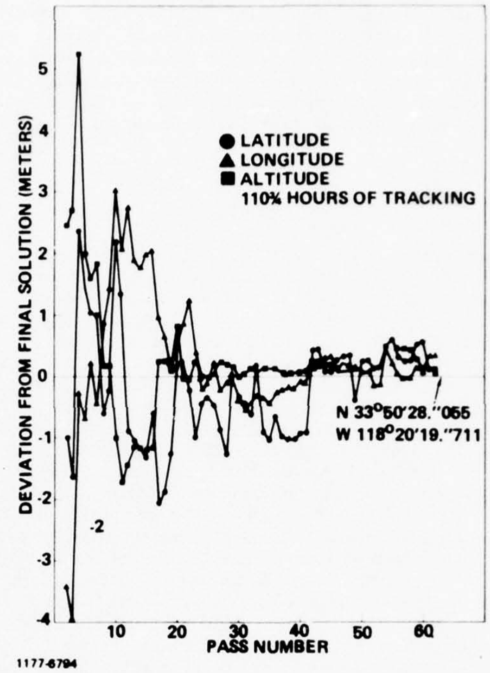


Figure 18. 3-D Point Positioning Convergence (62 MX 1502 Satellite Passes)

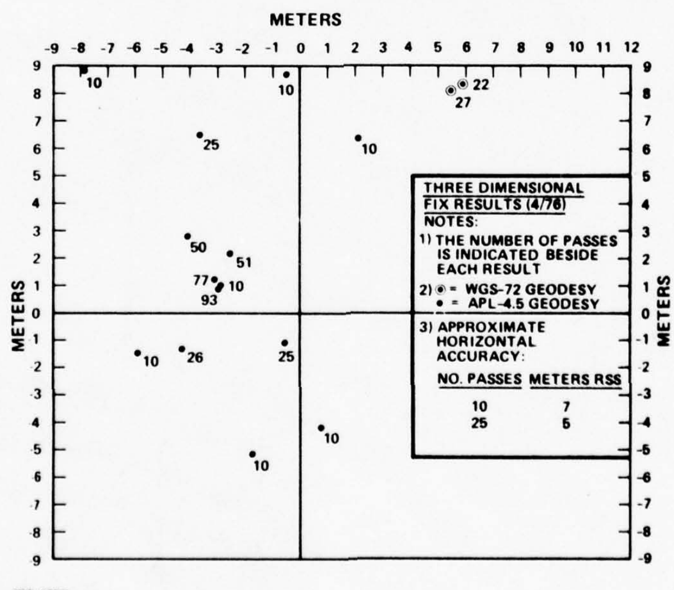


Figure 19. 3-D Point Positioning Results

Unfortunately, there is evidence that a precise ephemeris position fix result will differ from one using data from the satellite message. This difference is because the DMA uses a slightly different gravity model to compute satellite orbits than does the Navy Astronautics Group. This author regrets the difference and does not understand why it must persist.

3.5.5 Translocation Accuracy

Although precise ephemeris data are not available commercially, another technique called translocation can yield equivalent results. Advantage is taken of the fact

that almost all the error in a position fix is caused by factors external to the satellite receiver. Thus, two receivers tracking the same satellite pass at the same location should produce nearly the same result (i.e., the errors are strongly correlated). Experience has shown that the correlation is quite effective for interstation separations of 200 km or more. As a result, two or more stations can be located with respect to each other with an accuracy of 1 meter or better over very considerable distances.

One method of using translocation is to establish a base station which collects data from all available satellite passes for days or weeks. When fed to the 3-dimensional point positioning program, these data will yield an excellent absolute position determination. In the meantime, one or more portable receivers move from one location to another gathering 8 to 10 passes at each site. These data are then processed by translocation to define the position of each remote site with respect to the accurate base station location. An equally valid concept is to locate one station on an established and accepted geodetic reference point, thus using translocation to carry this geodetic reference to the remote sites.

Figures 20 and 21 show translocation results between two antennas which were very near each other so their relative position could be determined with great accuracy.

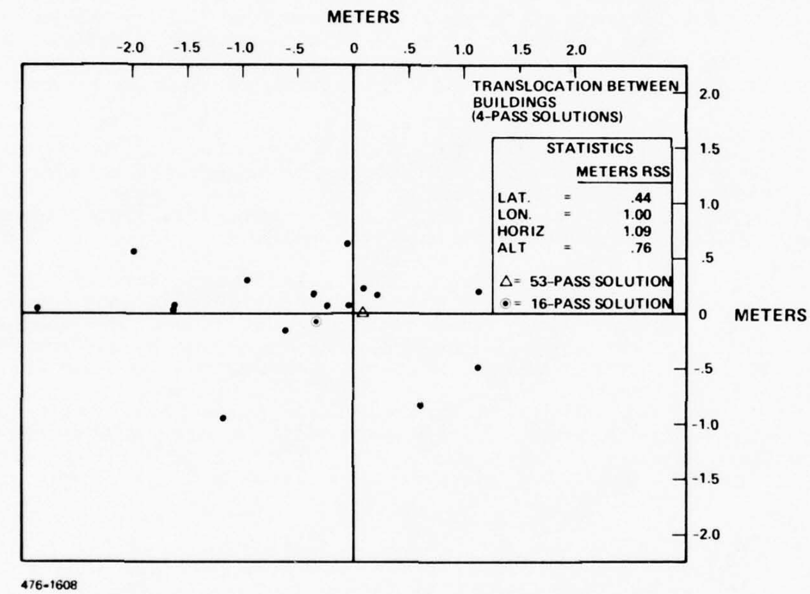


Figure 20. 4-Pass 3-D Translocation Results

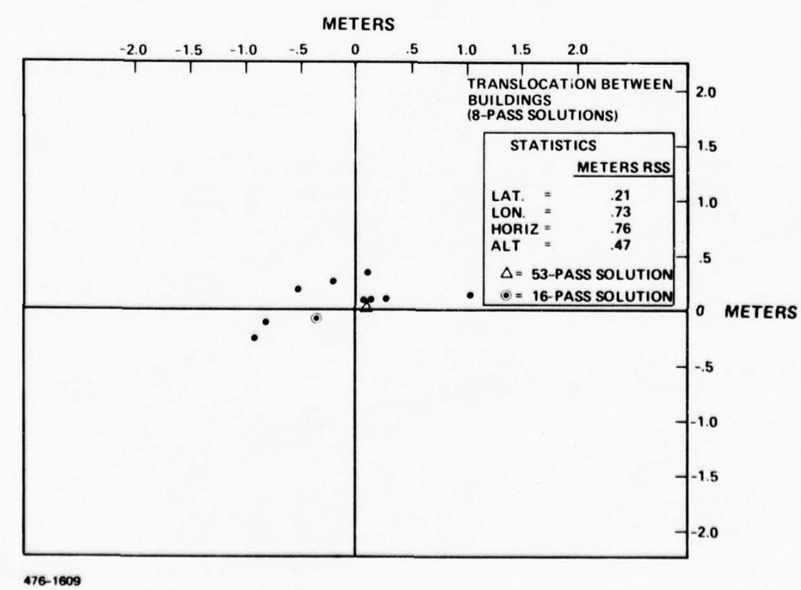


Figure 21. 8-Pass 3-D Translocation Results

Each dot shows the difference between the translocation result and the survey reference. All satellite passes above 15 degrees maximum elevation were used. For this test, manual editing forced a balance of east and west passes for the 4-pass solutions. For the 8-pass solutions, an imbalance of 5 vs. 3 was allowed. Otherwise, all other editing was performed automatically. The horizontal accuracy was 1.09 meters rms for the 4-pass solutions and 76 centimeters rms for the 8-pass solutions. This is a measure of quality both of the computer program and of the receivers being used for the test. It should be noted that slightly better results could be obtained through use of a rubidium or cesium frequency standard at each receiver. Field tests indicate that this level of translocation accuracy is obtainable over distances of several hundred kilometers.

3.6 Military Applications

The Transit system was developed initially to provide precise position updates for the Polaris submarine fleet. In this application, a submarine will expose its antenna at the appropriate time to update and to maintain the accuracy of its inertial navigation systems. Transit continues to be operated specifically to serve this Navy application.

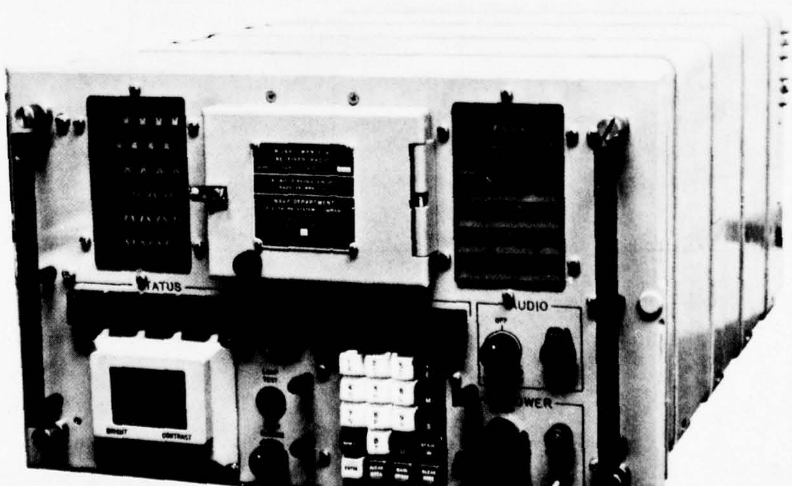
U.S. Navy attack submarines also are navigated by Transit. Figure 22 shows the AN/WRN-5 satellite navigator which was developed for use aboard nuclear attack submarines, although more are now being used aboard surface ships. A number of other Transit sets also are being used to navigate attack submarines, including the MX 702A/HP system shown in Figure 12 and, more recently, the MX 1102 Satellite Navigator shown in Figure 11. In fact, several NATO navies have expressed interest in a combination Transit-Omega navigator implemented within the MX 1102 structure both for submarines and for surface ships.

Submarine applications require the Transit navigator to provide satellite alert information so that appropriate times can be chosen to expose the antenna. In addition, it is desirable to minimize the duration of each antenna exposure. This requires a receiver such as the MX 1102 which tunes to the proper satellite frequency automatically, otherwise some provision for manual tuning must be provided.

Rather than tracking only selected satellite passes, surface ships track every available satellite pass. The navigation concepts, applications, and advantages are the same as for commercial ships, except that accurate, worldwide, all weather navigation also provides tactical and strategic advantages. Applications range from the navigation of major combat ships to patrol vessels guarding the 200 mile economic zone boundary.

Transit is used extensively for military land survey and mapping purposes. The U.S. Defense Mapping Agency and many of the NATO nations have cooperated on satellite survey operations across Europe. Equipment such as the AN/PPR-14 *Geoceiver*, shown in Figure 16, and the MX 1502 Satellite Surveyor, shown in Figure 17, can be used for these purposes.

As Transit user equipment has become smaller, more reliable, and less expensive, the opportunity for other land applications has been created. Magnavox is investigating the application of Transit fixes to vehicle positioning and even to manpack use. Although the time interval between Transit fixes is not desirable, there are many situations in which Transit could well be the only source of accurate geographic reference. This is particularily true for vast desert or jungle areas where accurately surveyed landmarks are not readily available.



1273-3581
UNCLASSIFIED

Figure 22. AN/WRN-5 Military Satellite Navigator

4.0 TRANSIT STATUS AND VITALITY

4.1 History and Future

Development of Transit began late in 1958, and the system became operational in January of 1964. On July 29, 1967, then Vice President Hubert H. Humphrey made an important announcement as part of a speech at Bowdoin College. The key paragraph from this speech reads as follows:

> "This week the President approved a recommendation that the Navy's Navigation Satellite System be made available for use by our civilian ships and that commercial manufacture of the required shipboard receivers be encouraged. This recommendation was developed by the Department of the Navy in support of initiatives of the Marine Sciences Council to strengthen worldwide navigational aids for civilian use. Our all-weather satellite system has been in use since 1964 by the Navy and has enabled fleet units to pinpoint their positions anywhere on the earth. The same degree of navigational accuracy will now be available to our non-military ships."

The use of Transit has expanded greatly in the years since its introduction. Manufacturers around the world have taken the Presidential encouragement literally, and since 1968 when the first commercial Transit sets were available, there has been a steady and dramatic increase in the types of equipment available and the number of users worldwide.

Regardless of past achievements, however, questions are raised about the future of Transit now that NAVSTAR, the Global Positioning System (GPS) is being developed. If GPS achieves its development objectives and operational funding is approved by the U.S. Congress, it is reasonable to expect that Transit will be discontinued after a sufficient overlap interval for users to depreciate existing equipment and to select appropriate replacement GPS equipment. Although no policy statement has been published at this time, the available information (see Reference 12) makes this author conclude that Transit will be available until at least 1995. The following paragraphs emphasize the vitality of the Transit system today and for the foreseeable future.

4.2 System Reliability and Availability

The Transit system reliability and availability can be seen in a number of areas. One is the remarkable success rate of the Navy Astronautics Group in maintaining a proper orbit message in the memory of each satellite. From January of 1964 to April of 1977, there had been only 7 message injections which were not verified as 100 percent successful out of a total 32,389 attempts. Each of the 7 was corrected on the next satellite pass, about 107 minutes later. This is a 99.98 percent success record and shows outstanding system reliability.

Figure 23 expresses the satellite status in terms of reliability and availability. Three of the five operational satellites were launched over ten years ago at this writing. Amazingly, the signals are strong and the satellites continue to function flawlessly. Backing up this group of "never say die" performers are twelve spacecraft stored where they were built many years ago at RCA Astro Electronics in New Jersey.

Being very light (about 61 kilograms), Transit satellites can be placed in their 1,100 kilometer orbits with relatively inexpensive, solid fuel Scout rockets. Nine of these boosters currently are in reserve to support future launches.

It appears that Transit is in extremely good health when it comes to reliable performance today and provision for continuation of service for many years to come, especially noting the proven longevity of the spacecraft design.

4.3 New Generation of Satellites

As shown in Figure 24, the Applied Physics Laboratory has developed a new generation of Transit satellites, which they called TIP for Transit Improvement Program. Two prototype satellites were launched as part of the development effort.

The Navy has decided to produce a limited number of these new satellites, which will now be called NOVA. RCA is building the first three NOVA satellites, and it is expected that at least two more will be built. The first NOVA is expected to be launched in the third quarter of 1979. This new satellite will be especially welcome in filling the orbit gap now existing between satellites 30120 and 30200, as discussed in Section 4.7.

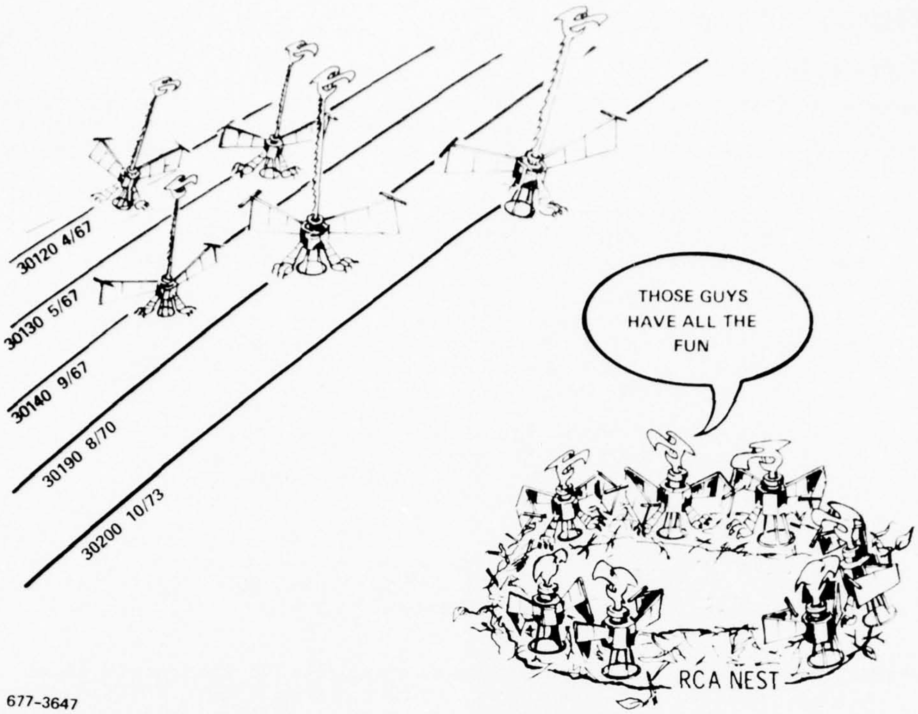


Figure 23. The 5 Operational Transit Satellites, Launched on the Dates Shown, are Backed by Twelve Reserve Spacecraft at RCA

The NOVA satellite signals are entirely compatible with the existing Transit satellite signals. Therefore, all users will have access to this new spacecraft. However, the NOVA satellites provide many important new capabilities, all of which have been verified with the experimental TIP satellites. Of particular interest are the following:

- DISCOS, for disturbance compensation system, eliminates the effect of atmospheric drag. As a result, each orbit determination will retain accuracy for up to a week instead of 24 hours now. With NOVA, we expect survey navigation results to converge faster and have better accuracy.
- NOVA is controlled by an onboard general purpose digital computer which can be programmed from the ground. In conjunction with a larger memory, the computer can provide orbit parameters for ten days without requiring upload of new information.
- A new data modulation, transparent to existing receivers, can be switched on. Plans for this modulation have not been announced, but it could be used to provide more precise orbit parameters.
- The received signal level from NOVA satellites will be twice as strong (3 dB). Antenna polarization will be left hand circular on both channels rather than left on 150 MHz and right on 400 MHz at present.

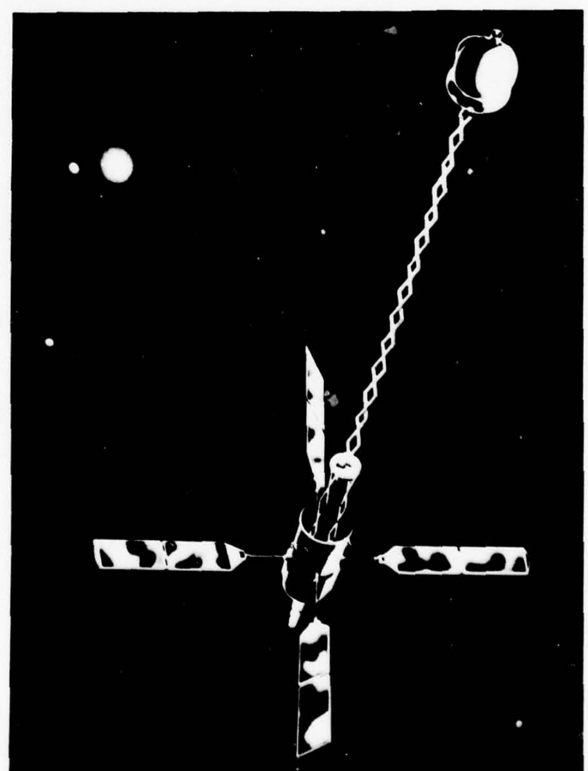


Figure 24. New Generation NOVA Transit Satellite (Previously called TIPS)

- Very precise clock control has been achieved by permitting the onboard computer to adjust oscillator frequency with a resolution of about 1×10^{-12}. (To make the carrier and the data modulation coherent, the nominal frequency offset has been changed from 80 ppm to 84.48 ppm, which should not cause compatibility problems.)

- To transmit the precise time information, a high frequency pseudo-random noise (PRN) modulation has been added to both the 150 and 400 MHz signals. This also can be used to achieve single-channel, refraction corrected fixes (by detecting the difference in group delay and phase delay effects) and a properly equipped receiver can block out signals from any other satellite, thus eliminating the potential for cross-satellite interference.

4.4 Expanding User Base

Figure 25 is a chart prepared by the Navy Astronautics Group based on information received from 15 of 19 manufacturers of Transit user equipment. The chart shows a total user population of 1,899 sets at the beginning of 1977, which was expected to grow to 4,350 sets by the end of 1978.

The user population growth predicted by the manufacturers represents an annual growth rate of 51 percent. To see if this were possible, data was included from an earlier survey showing the total population at the beginning of 1974 to be 600 sets. Growing from 600 to 1,899 in three years required an annual rate of 47 percent. Thus, the predicted annual growth of 51 percent appears to be in line with past trends, and it may be conservative when recent product innovations are considered.

Figure 25 shows the growth as a linear function of time, but including the data from 1974 tells us that this is not the case. In fact, the number of users has been increasing as a percentage of the existing population, which is a straight line on logarithmic paper. Figure 26 is such a plot using the three data points provided by the Navy Astronautics Group. What may be surprising is that at present rates the user population should reach 10,000 by the early 1980's. Based on data available as of the first quarter of 1978, this growth trend appears to be continuing.

4.5 Investment in Transit Navigation Equipment

Combining data from the Navy Astronautics Group with other sources, the total investment in Transit navigation equipment has been estimated, as summarized by Figure 27. Research and development costs are not included, and equipment known to be out of service has been deleted. Overall, we believe the estimates are on the low side.

The Navy Strategic Systems Project Office has been included as a separate category due to their special involvement with Transit. The total U.S. Government investment in Transit user equipment is nearly 45 million dollars. Most of the integrated systems are owned and operated by private firms engaged in offshore oil exploration. The remaining dual-channel navigation systems are used for survey work of various types, such as oceanography, land survey, drill rig positioning, cable laying, etc. The single-channel navigators are used for general navigation purposes where 0.1 mile fix accuracy is sufficient, and this is the area of fastest growth.

The last column in Figure 27 is an estimate of the cost of equipment plus spares. Ten percent spares cost was assumed for the single-channel equipment and 30 percent for all other categories.

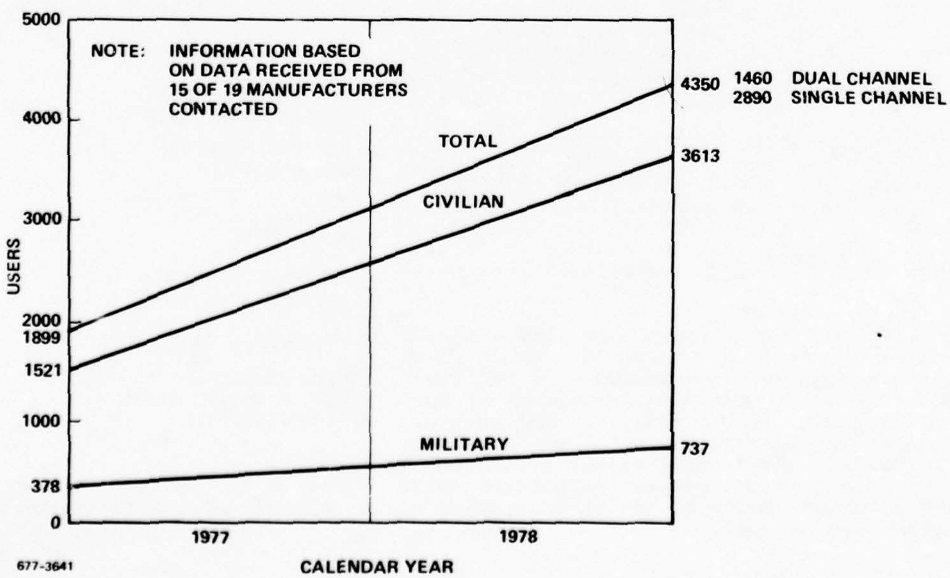
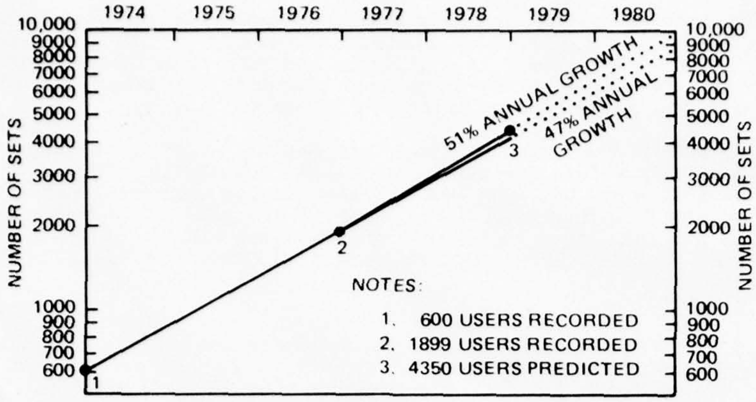


Figure 25. Present Status and Expected Growth in Number of Transit System Users (Provided by the Navy Astronautics Group)



677-3642

Figure 26. Growth of Transit User Population Obtained From Data Provided by the Navy Astronautics Group

CATEGORY	QUANTITY	AV. COST (THOUSANDS)	TOTAL COST (MILLIONS)	WITH SPARES (MILLIONS)
NAVY STRATEGIC SYSTEMS PROJECT OFFICE	73	$ 251	$ 18.4	$ 23.9
U.S. GOVERNMENT – ALL OTHER	469	56	26.3	34.2
INTEGRATED SYSTEMS	118	231	27.3	35.5
OTHER DUAL-CHANNEL	539	47	25.2	32.8
SINGLE-CHANNEL	2239	22	48.4	53.2
TOTALS	3438		$145.6	$179.6

478-1513

Figure 27. Estimated Investment in Transit Navigation Equipment (April 1978)

Figures 26 and 27 carry a powerful and surprising message. It is probable that at this time more money has been invested in Transit user equipment than in marine equipment for any other U.S. radionavigation system, including Loran-A, Loran-C, or Omega. Naturally the reason for this has been the much higher price for Transit equipment, which always requires a computer and often is combined with other sensors to form an integrated system. However, Figure 26 shows that the user population also is growing rapidly, spurred by technical innovations which permit lower prices, better performance, and greater reliability.

4.6 Cost of Transit System Operation

The cost of operating Transit has been estimated by the Navy to be as shown in Figure 28. For those familiar with the operational costs of any other major navigation system, it should be obvious that Transit is very inexpensive to operate and to maintain.

4.7 Improvement in Orbital Coverage

Figure 29 shows the orbital spacing of the five operational Transit satellites and their rates of precession as of March 23, 1978. This specific orbital configuration was used to predict the average interval between satellite fixes given by Figure 3.

A better way to visualize the interval between fixes is that of Figure 30, which shows the cumulative waiting time probability at three different latitudes. Note that intervals of more than 12 hours occur infrequently at the equator, and intervals of six to seven hours occur at higher latitudes. These peak values are strongly related to the large gap between satellites 30120 and 30200 shown in Figure 29, which is growing at about 5.1 degrees per year.

TRANSIT GROUND STATION	PERSONNEL
POINT MUGU, CALIFORNIA	152
PROSPECT HARBOR, MAINE	20
ROSEMONT, MINNESOTA	28
WAHIAWA, HAWAII	9
TOTAL	209
ANNUAL SUPPORT	**ANNUAL COST**
TRANSIT GROUP SUPPORT	$ 5.0 M
STORAGE OF 12 SATELLITES	0.3
SATELLITE REPLACEMENT COST	**EACH**
(INCLUDES SCOUT LAUNCH VEHICLE, SATELLITE CHECKOUT, AND LAUNCH SUPPORT.)	$ 3.5 M

478-1517

Figure 28. Cost of Operating the Transit System (Provided by the U.S. Navy, April 1977)

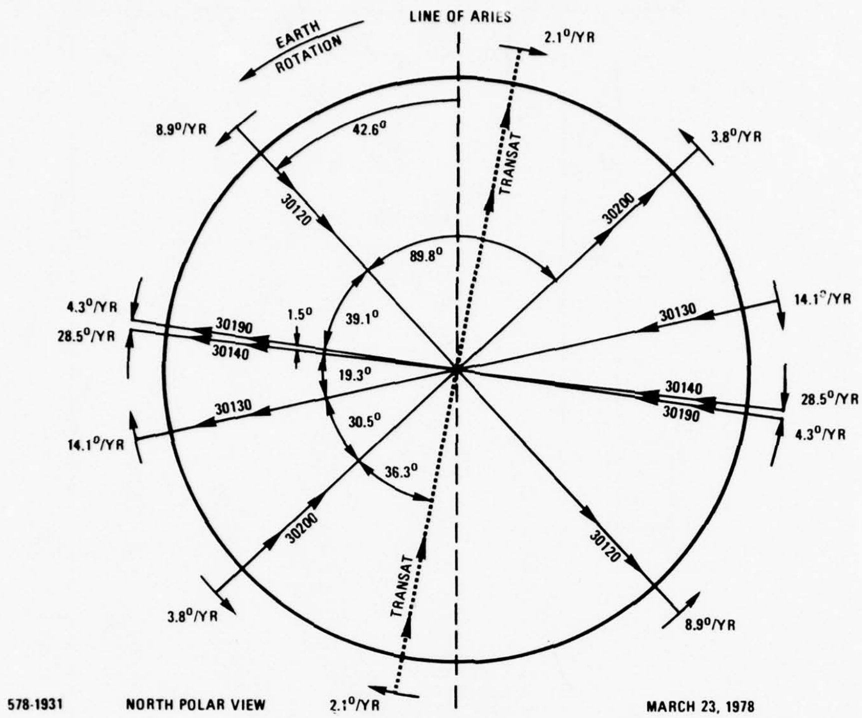


Figure 29. Orbital Separation of the Five Operational Transit Satellites and TRANSAT (30110) on March 23, 1978

To evaluate the effect of filling the gap with another satellite, the interval prediction program also was run with six satellites. The sixth satellite is TRANSAT (30110), shown with a dotted line in Figure 29, which was launched by the U.S. Navy in 1977. This satellite is intended for purposes other than navigation, although it has a Transit navigation mode which can switched on if desired.

Figure 31, when compared with Figure 30, shows the dramatic effect of having a satellite in the orbit coverage gap. Not only are there more satellite fixes available, but a much higher percentage occur after shorter waiting times. Figure 32 shows the effect on mean time between fixes of having TRANSAT.

Although having the gap filled would be very desirable, the Navy does not plan to use TRANSAT in this way. However, as described in Section 4.3, the Navy does plan to launch the new generation of NOVA satellites beginning in the third quarter of 1979. Not only will NOVA fill the gap, but the orbits will be controlled to maintain precession at negligible levels. In 1980, two NOVA satellites with orthogonal orbits will form the backbone of the Transit system, with the existing satellites continuing to provide fixes as well.

4.8 Summary

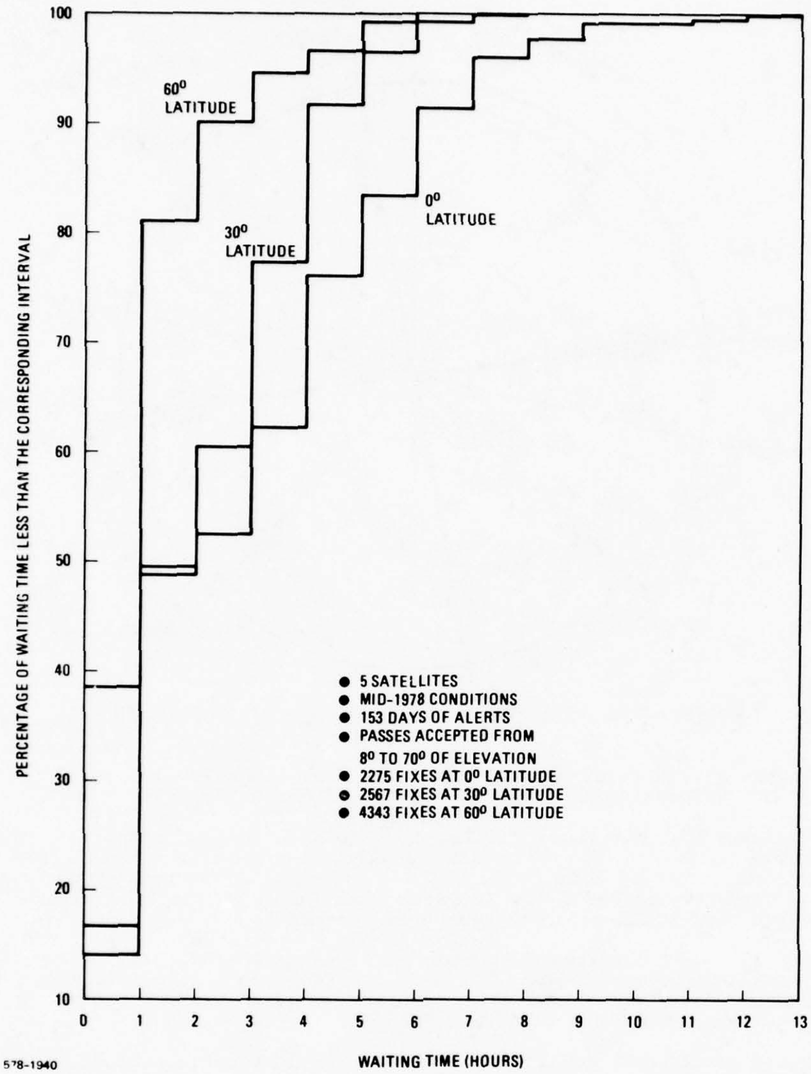
The preceding paragraphs have attempted to communicate the basic vitality of the Transit system. We see this vitality in the system reliability, the new generation of satellites, the expanding user base, the amazing breadth of applications, the substantial worldwide investment in Transit navigation equipment, and in the very low cost of system operation. With all things considered, this author is certain the Transit system will continue to provide its vital navigation service until at least 1995.

5.0 THE POSITION FIX TECHNIQUE

5.1 The Satellite Signals

Figure 33 is a block diagram of the Transit satellite electronics. The satellites transmit coherent carrier frequencies at approximately 150 and 400 MHz. Because both signals are derived by direct multiplication of the reference oscillator output, the transmitted frequencies are very stable, changing no more than about 1 part in 10^{11} during a satellite pass. Thus, they may be assumed to be constant with negligible error.

The reference oscillator output also is divided in frequency to drive the memory system. In this way, the navigation message stored there is read out and encoded by phase modulation onto both the 150 and 400 MHz signals at a constant and carefully



578-1940

Figure 30. Cumulative Probability of Waiting Time for the Next Transit Fix With the Five Current Satellites (mid-1978)

controlled rate. Thus, the transmitted signals provide not only a constant reference frequency and a navigation message but also timing signals, because the navigation message is controlled to begin and to end at the instant of every even minute. An updated navigation message and time corrections are obtained periodically from the ground by way of the satellite's injection receiver. The time correction data are stored in the memory and applied in steps of 9.6 microseconds each.

Each binary bit of the message is transmitted by phase modulation of the 150 and 400 MHz signals. The modulation format for a binary one is given in Figure 34, and a binary zero is transmitted with the inverse pattern. As shown, this format furnishes a clock signal at twice the bit rate, which is used to synchronize the receiving equipment with the message data.

Because the satellites transmit only about one watt of power and may be thousands of kilometers away, very sensitive receivers are needed. In addition, however, the orbit parameters must be verified by comparing redundant messages to detect and eliminate occasional errors in the received data.

5.2 Interpretation of Satellite Message

Figure 35 indicates how the navigation message defines the position of the satellite. During every two minute interval the satellite transmits a message consisting of 6,103 binary bits of data organized into 6 columns and 26 lines of 39-bit words, plus a final 19 bits. The message begins and ends at the instant of the even minute, which

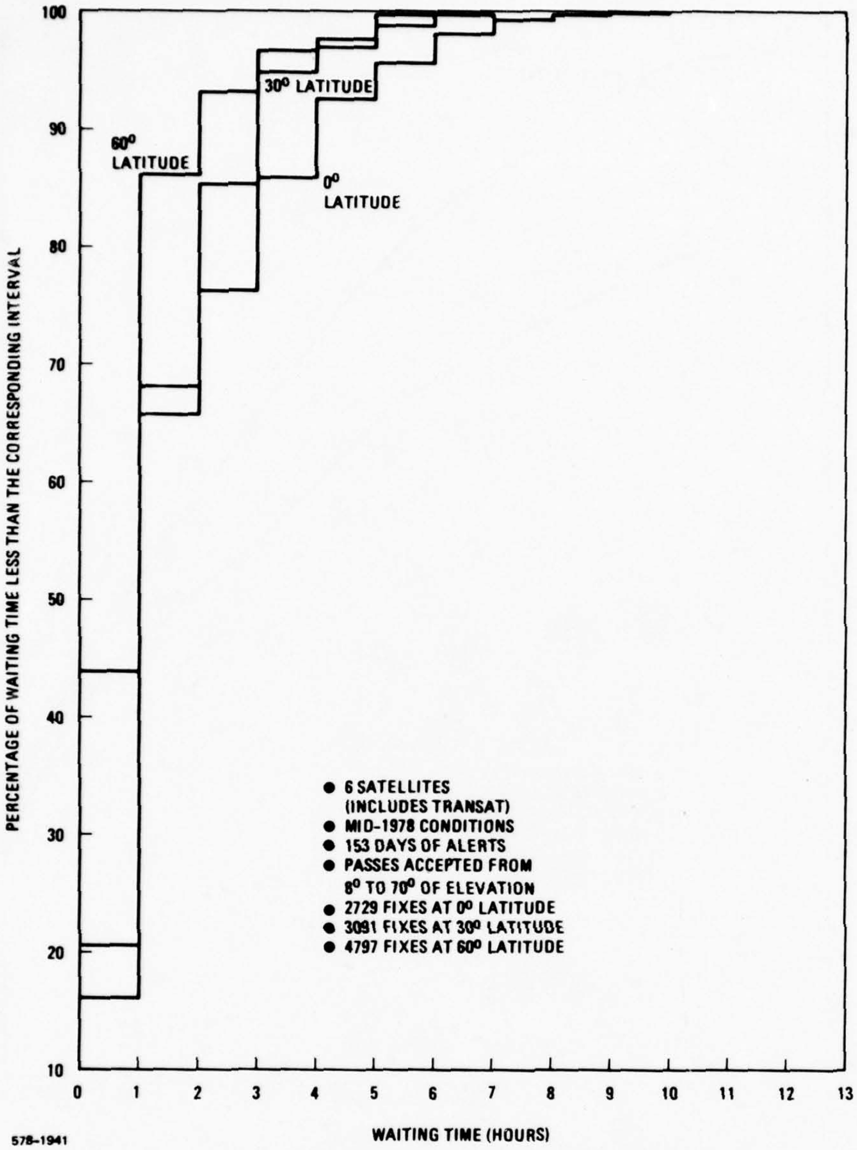


Figure 31. Cumulative Probability of Waiting Time for the Next Transit Fix Assuming TRANSAT Use (mid-1978)

are denoted as time marks t_i and t_{i+1}. The final 25 bits of each message form a synchronization word (0111111111111111111111110) that identifies the time mark and the start of the next 2-minute message. By recognizing this word, the navigation receiver establishes time synchronization and thereafter can identify specific message words.

The orbital parameters are located in the first 22 words of column 6, and those in lines 9 through 22 are changed only when a new message is injected into the memory. These fixed parameters define a smooth, precessing, elliptical orbit; satellite position being a function of time since a recent time of orbit perigee.

The words in lines one through eight shift upward one place every two minutes, with a new word inserted each time in line eight. These variable parameters describe the deviation from the smooth ellipse of the actual satellite position at the indicated even minute time marks. By interpolation through the individual variable parameters, the satellite position can be defined at any time during the satellite pass.

Figure 36 aids in interpreting the Transit message parameters. On the left is a set of typical fixed parameters and an indication of how they are to be interpreted. On the right is a set of variable parameters with an interpretation of one. The following paragraphs will describe how each of these is used.

The "Q" number provides a time tag for each word of variable parameters. In the example given, the number 07 means that this word applies at seven 2-minute intervals past the half hour, i.e., 14 minutes or 44 minutes after the hour. This is why it is necessary to initialize a Transit set to within plus or minus 15 minutes of correct (GMT) time in order to synchronize properly. A time error of less than 15 minutes will be resolved by the "Q" number from the satellite message.

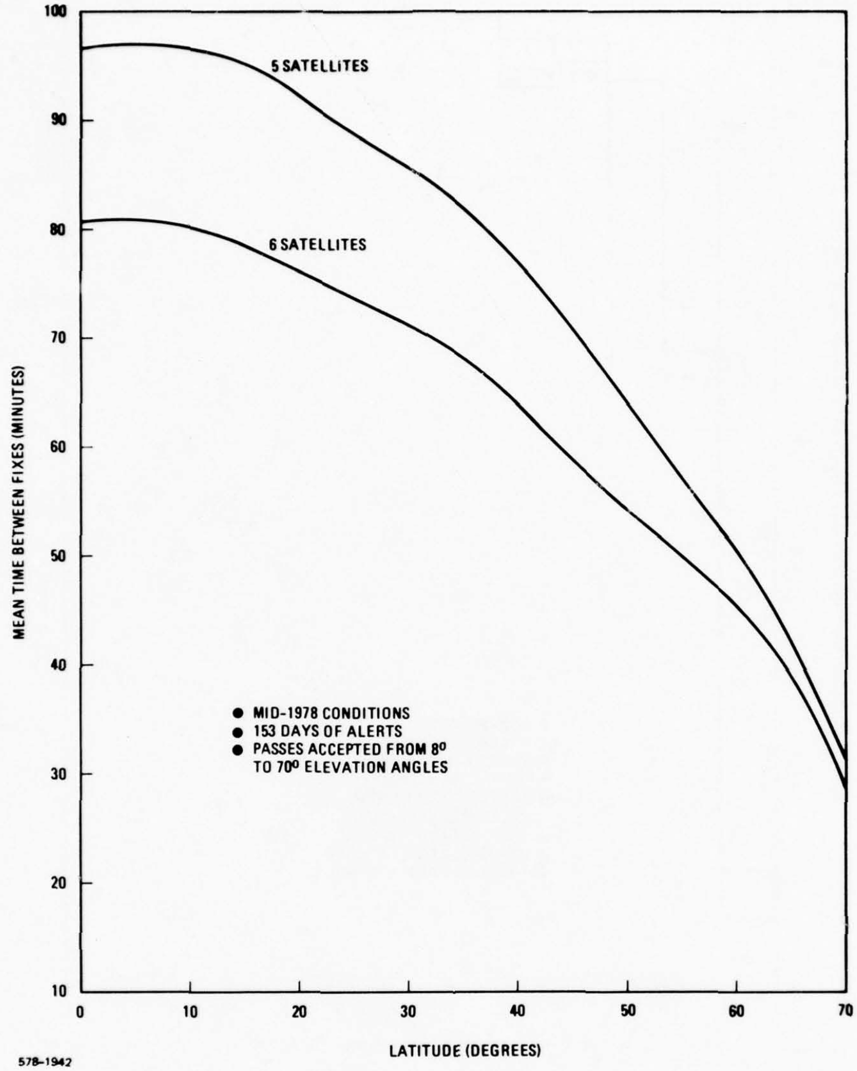


Figure 32. Mean Time Between Fixes Which Would Occur With and Without TRANSAT During mid-1978

From Figure 36 also note that only one digit of the variable parameter η_k is transmitted in each word. Because two digits are required, this parameter is defined *only every four minutes at times which are integer multiples of four minutes.* The interpretation of the first digit of η_k also is given by the figure.

The objective is to define the satellite position as a function of time. To achieve this, three different coordinate systems are employed. Figure 37 defines the u, v, w coordinate system. These coordinates are earth-centered, u and v lie in the plane of the satellite orbit, and u is through the point of perigee (closest point to the earth). On the left of Figure 37 are shown the classical Kepler orbit definition equations. The Transit orbit definition equations are very similar, except for simplifications in the expressions for E_k and for v_k. Error introduced by these simplifications is eliminated by application of variable parameters ΔE_k and ΔA_k. The w_k parameter defines out-of-plane satellite motion, which is simply the variable parameter η_k.

Figure 38 shows how the x', y', z' coordinates are obtained by rotation of the u, v, w coordinates. Rotation by the "argument of perigee" places x' in the earth's equatorial plane.

Finally, Figure 39 shows that with two rotations the satellite position can be defined in an X, Y, Z coordinate system which is earth-centered and earth fixed, with Z being the polar axis (mean pole of 1900-1905 or Conventional International Origin) and X being in the equatorial plane through the Greenwich meridian. The two rotations account for the longitude of the orbit plane at t_k and the inclination of the orbit with respect to earth's equatorial plane.

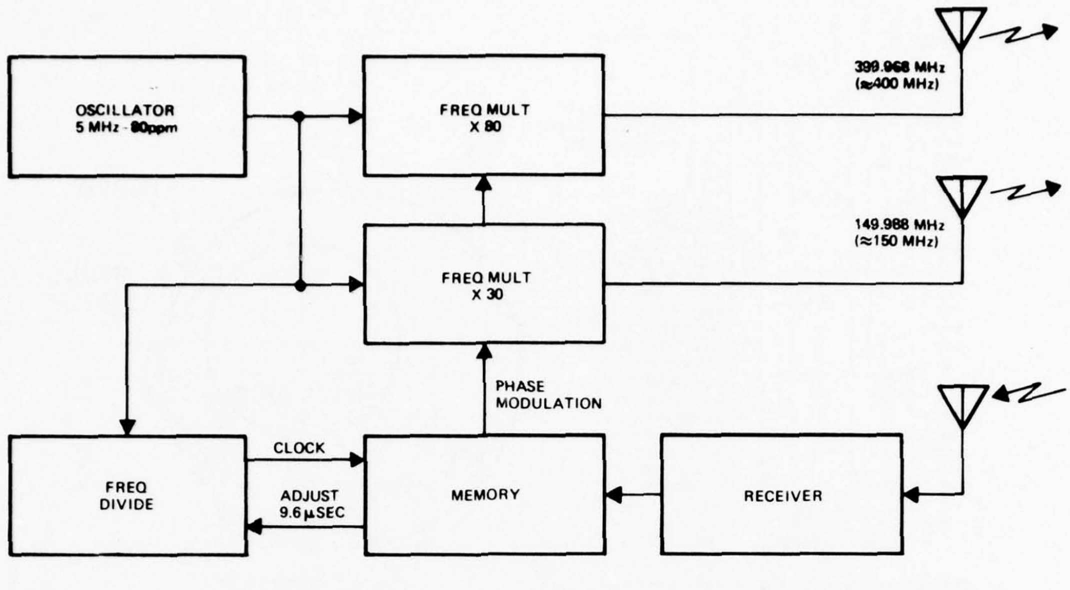


Figure 33. Transit Satellite Block Diagram

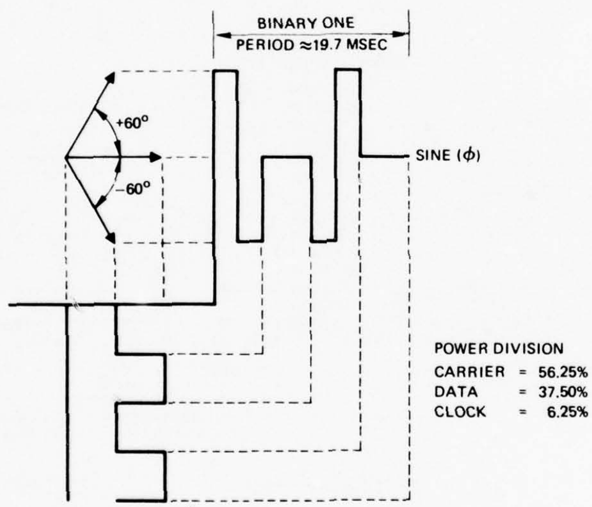


Figure 34. Transit Data Phase Modulation

Figures 37 through 39 clearly show how the Transit orbit parameters are interpreted and how they are used to obtain a definition of the satellite position in an earth fixed Cartesian coordinate frame. These coordinates can be computed for any time by interpolation of the satellite variable parameters.

5.3 The Doppler Measurement

By receiving the navigation message the Transit system user learns the position of the satellite as a function of time. Thus, to obtain a position fix he must relate his position to the known satellite orbit. This relationship is established by measuring the Doppler shift, which is a unique function of the observer's position and motion relative to the known satellite orbit.

Figure 40 illustrates the Doppler measurement technique. The frequency f_R being received from the satellite consists of the frequency being transmitted f_T plus a Doppler frequency shift of up to ±8 kHz due to relative motion between the satellite and the receiver. Note that the transmitted frequency is offset low by about 80 ppm (32 kHz at 400 MHz) to prevent f_R from crossing 400 MHz.

DATUM	SPHEROID	SEMI-MAJOR AXIS	RECIPROCAL FLATTENING	SHIFT TO WGS-72 a = 6378135 1/f = 298.26

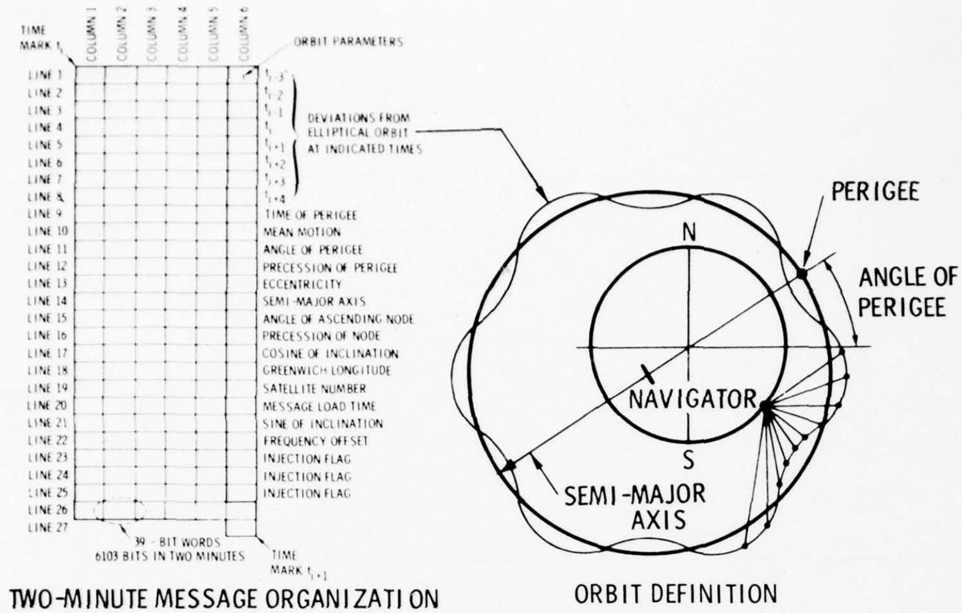


Figure 35. Satellite Message Describes Orbital Position

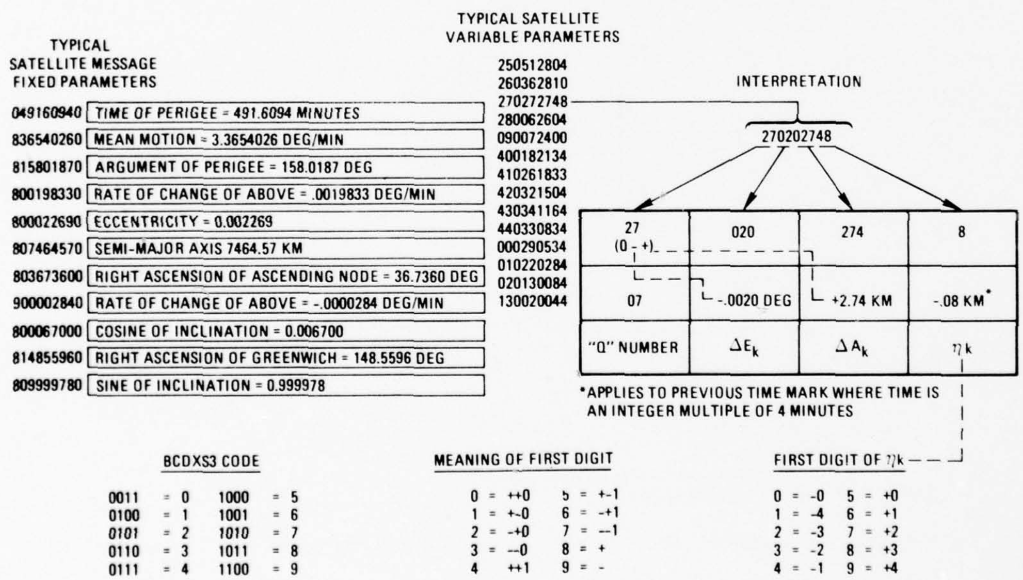


Figure 36. Interpretation of the Transit Message Parameters

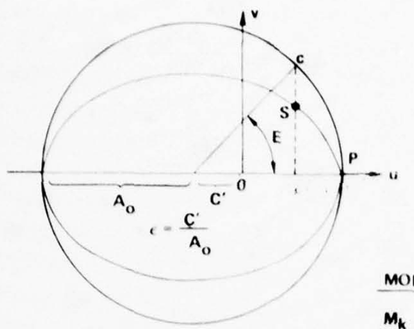
The navigation receiver is equipped with a stable reference oscillator from which a 400 MHz ground reference frequency f_G is derived. Oscillator stability must be adequate to assume a constant frequency throughout the satellite pass. As shown by the figure, the navigation receiver forms the difference frequency f_G-f_R, and each Doppler measurement is a count of the number of difference frequency cycles occuring between time marks received from the satellite. Because every message bit effectively represents another time mark, the Doppler counting intervals are formed with respect to the message format of Figure 35. For example, each line of the message lasts about 4.6 seconds, and the commonly used Doppler count interval of 23 seconds is formed by starting a new count at the end of every fifth line.

Each Doppler count is composed of two parts: the count of a constant difference frequency f_G-f_T, minus the count of the number of Doppler cycles received during

2-40

A THE STANDARD MOLODENSKY FORMULAS

N = GEOID-ELLIPSOID SEPARATION. THE DISTANCE OF THE GEOID ABOVE (+N) OR BELOW (-N) THE ELLIPSOID



CLASSICAL ORBIT DEFINITION

$M(t) = n(t - t_p)$

$E(t) = M(t) + \epsilon \text{ SIN } E(t)$

$A = A_o$

$u(t) = A(\text{COS } E(t) - \epsilon)$

$v(t) = A\sqrt{1-\epsilon^2} \text{ SIN } E(t)$

w(t) IS UNDEFINED

0 = EARTH CENTER
S = SATELLITE POSITION
P = PERIGEE
M = MEAN ANOMALY
E = ECCENTRIC ANOMALY
A_o = SEMI-MAJOR AXIS

MODIFIED TRANSIT ORBIT DEFINITION

$M_k = n(t_k - t_p)$

$E_k = M_k + \epsilon \text{ SIN } M_k + \Delta E(t_k)$

$A_k = A_o + \Delta A(t_k)$

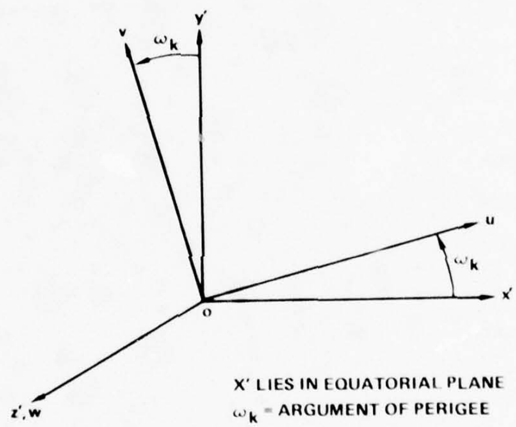
$u_k = A_k(\text{COS }(E_k) - \epsilon)$

$v_k = A_k \text{ SIN }(E_k)$

$w_k = \eta(t_k)$

877-5083

Figure 37. u, v, w Satellite Coordinates are Earth-Centered and Aligned With Perigee



$\omega_k = \omega_o - \dot{\omega}\Delta t_K$

$X'_k = U_k \text{ COS } \omega_k - V_k \text{ SIN } \omega_k$

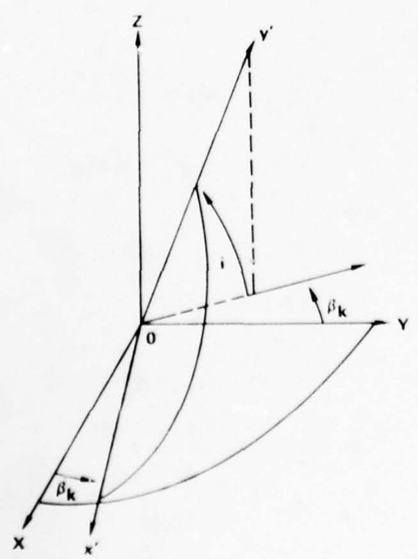
$Y'_k = U_k \text{ SIN } \omega_k + V_k \text{ COS } \omega_k$

$z'_k = \eta_k$

X' LIES IN EQUATORIAL PLANE
ω_k = ARGUMENT OF PERIGEE

877-5084

Figure 38. x', y', z' Satellite Coordinates are Earth-Centered With x' in the Equatorial Plane



Z = POLAR AXIS
X = THROUGH GREENWICH MERIDIAN
i = ORBIT INCLINATION
β_k = LONGITUDE OF ORBIT PLANE
$\beta_k = (\Omega_o - \Lambda_g) + (\dot{\Omega} - \omega_e)\Delta t_k$
$X_{sk} = x'_k \text{ COS } \beta_k - y'_k C_i \text{ SIN } \beta_k + z'_k S_i \text{ SIN } \beta_k$
$Y_{sk} = x'_k \text{ SIN } \beta_k + y'_k C_i \text{ COS } \beta_k - z'k S_i \text{ COS } \beta_k$
$Z_{sk} = y'_k S_i + z'_k C_i$

877-5086

Figure 39. X, Y, Z Satellite Coordinates are Earth-Centered and Earth Fixed

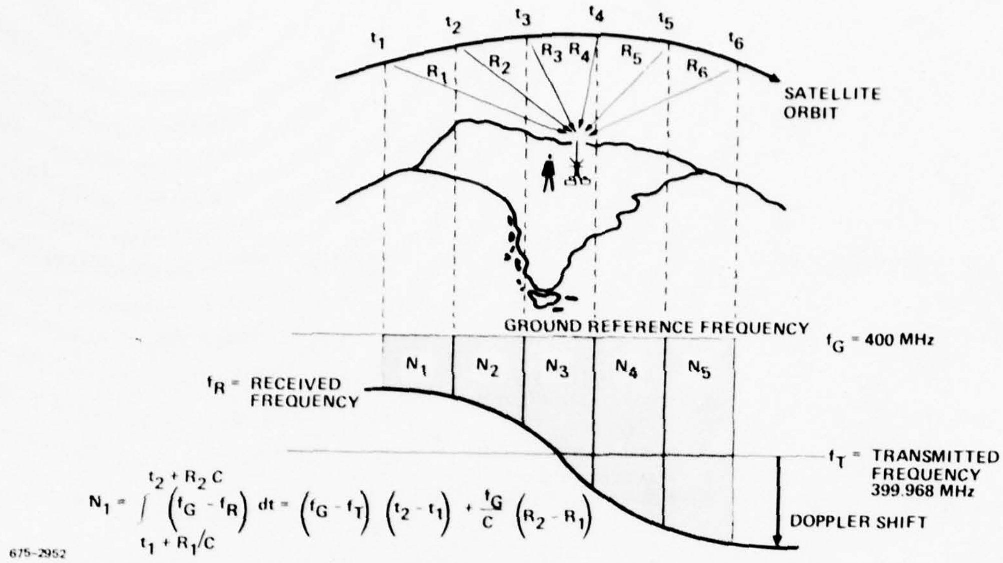


Figure 40. Each Doppler Count Measures Slant Range Change

that time interval. It is the Doppler cycle count which is physically meaningful. The count of the difference frequency is an additive constant which is eliminated by the position fix calculation.

Figure 40 emphasizes that the distance between the satellite and the observer changes throughout the satellite pass. It is this change, in fact, which causes the Doppler frequency shift. As the satellite moves closer, more cycles per second must be received than were transmitted to account for the shrinking number of wavelengths along the propagation path. For each wavelength the satellite moves closer, one additional cycle must be received. Therefore, the Doppler frequency count is a direct measure of the change in distance between the receiver and the satellite over the Doppler count interval. In other words, the Doppler count is a geometric measure of the range difference between the observer and the satellite at two points in space, accurately defined by the navigation message. This is a very sensitive measure because each count represents one wavelength, which at 400 MHz is only 0.75 meter.

The equation defining the Doppler count of $f_G - f_R$ is the integral of this difference frequency over the time interval between receipt of time marks from the satellite. For example,

$$N_1 = \int_{t_1 + R_1/C}^{t_2 + R_2/C} (f_G - f_R)\, dt \qquad (1)$$

Note that $t_1 + R_1/C$ is the time of receipt of the satellite time mark that was transmitted at time t_1. The signal is received after propagating over distance R_1 at the velocity of light C.

Equation 1 represents the actual measurement made by the satellite receiver, but it is helpful to expand this equation into two parts:

$$N_1 = \int_{t_1 + R_1/C}^{t_2 + R_2/C} f_G\, dt - \int_{t_1 + R_1/C}^{t_2 + R_2/C} f_R\, dt \qquad (2)$$

Because the first integral in Equation 2 is of a constant frequency f_G, it is easy to integrate, but the second integral is of the changing frequency f_R. However, the second integral represents the number of cycles <u>received</u> between the times of <u>receipt</u> of two timing marks. By a "conservation of cycles" argument, this quantity must <u>equal identically</u> the number of cycles <u>transmitted</u> during the time interval between <u>transmission</u> of these time marks. Using <u>this identity</u>, Equation 2 can be written

$$N_1 = \int_{t_1 + R_1/C}^{t_2 + R_2/C} f_G\, dt - \int_{t_1}^{t_2} f_T\, dt \qquad (3)$$

Because the frequencies f_G and f_T are assumed constant during a satellite pass, the integrals in Equation 3 become trivial, resulting in

$$N_1 = f_G\left[(t_2 - t_1) + \frac{1}{C}(R_2 - R_1)\right] - f_T(t_2 - t_1) \tag{4}$$

Rearranging the terms in Equation 4 gives

$$N_1 = (f_G - f_T)(t_2 - t_1) + \frac{f_G}{C}(R_2 - R_1) \tag{5}$$

Equation 5 clearly shows the two parts of the Doppler count. First is the constant difference frequency multiplied by a time interval defined by the satellite clock. Second is the direct measure of slant range change measured in wavelengths of the ground reference frequency C/f_G. It happens that the wavelength of f_G is the proper scale factor because received time marks are used to start and stop the Doppler counts. If a ground clock is used to control the count intervals, the wavelength of f_T would become the appropriate scale factor.

5.4 Computing the Fix

A usable satellite pass will be above the horizon between 10 and 18 minutes, which determines the number of Doppler counts acquired. Typically 20 to 40 counts will be collected by modern equipment. The Doppler counts and the satellite navigation message are passed to a small digital computer for processing. For simplicity, we will assume a stationary receiver as shown in Figure 40 in order to establish the basic position fix concept.

The computer first takes advantage of message redundancy to eliminate errors in the received orbit parameters. It is then able to compute the satellite's position at the beginning and end of every Doppler count. The computer also receives an estimate of the navigator's position in three dimensions, i.e., latitude, longitude, and altitude above the reference ellipsoid. The equations of Figure 41 are used to convert the navigator's position into the same Cartesian coordinate system shown in Figure 39, which permits the slant range from the navigator to each satellite position to be calculated. It is then possible to compare the slant range change measured by each Doppler count with the corresponding value computed from the estimated navigator's position.

The difference between a Doppler measured slant range change and the value computed from the estimated position is called a residual e_i. The objective of the position fix calculation is to find the navigator's position which minimizes the sum of the squares of the residuals (i.e., makes the calculated slant range change values agree best with the measured values). To implement the solution, a simple, linear estimate is made of the effect each variable will have on each residual. Assuming we wish to solve for latitude (ϕ), longitude (λ), and the unknown frequency offset $F = f_G - f_T$, we can write

$$\hat{e}_i = e_i - \frac{\partial e_i}{\partial \phi}\Delta\phi - \frac{\partial e_i}{\partial \lambda}\Delta\lambda - \frac{\partial e_i}{\partial F}\Delta F \tag{6}$$

This equation states that if we move the estimated position by $\Delta\phi$ and by $\Delta\lambda$ and the estimated frequency offset by ΔF, the present residual e_i will become a new value, estimated to be $\hat{e}_i$. Next we wish to minimize the sum of the squares of the estimated residuals by setting the partial derivative with respect to each variable equal to zero. This results in three equations, where the summation covers the m valid Doppler count residuals.

$$\begin{aligned}
\frac{\partial}{\partial \phi}\sum_{i=1}^{m}\hat{e}_i^2 &= 2\sum_{i=1}^{m}\left(\hat{e}_i \cdot \frac{\partial \hat{e}_i}{\partial \phi}\right) = 0 \\
\frac{\partial}{\partial \lambda}\sum_{i=1}^{m}\hat{e}_i^2 &= 2\sum_{i=1}^{m}\left(\hat{e}_i \cdot \frac{\partial \hat{e}_i}{\partial \lambda}\right) = 0 \\
\frac{\partial}{\partial F}\sum_{i=1}^{m}\hat{e}_i^2 &= 2\sum_{i=1}^{m}\left(\hat{e}_i \cdot \frac{\partial \hat{e}_i}{\partial F}\right) = 0
\end{aligned} \tag{7}$$

ϕ	= LATITUDE	
λ	= LONGITUDE	WGS-72 VALUES
H	= HEIGHT ABOVE ELLIPSOID	
a	= SEMI-MAJOR AXIS	(6378135 METERS)
f	= FLATTENING COEFFICIENT	(1/298.26)
b	= a (1-f) = SEMI-MINOR AXIS	(6356750.52 METERS)
e	= $\sqrt{f(2-f)}$ = ECCENTRICITY	
RN	= RADIUS OF CURVATURE IN THE PRIME VERTICAL	
RN	= $a/(1-e^2 \, SIN^2\phi)^{1/2}$	
XN	= $(RN + H) \, COS\,\phi \, COS\,\lambda$	
YN	= $(RN + H) \, COS\,\phi \, SIN\,\lambda$	
ZN	= $[RN\,(1-e^2) + H] \, SIN\,\phi$	

476-1610

Figure 41. Relating Latitude and Longitude to Cartesian Coordinates

Ignoring all but the first-order terms of Equations 7 gives three equations in the three selected variables, $\Delta\phi$, $\Delta\lambda$, and ΔF.

$$\sum_{i=1}^{m} \frac{\partial e_i}{\partial \phi} \left[e_i - \frac{\partial e_i}{\partial \phi} \Delta\phi - \frac{\partial e_i}{\partial \lambda} \Delta\lambda - \frac{\partial e_i}{\partial F} \Delta F \right] = 0$$

$$\sum_{i=1}^{m} \frac{\partial e_i}{\partial \lambda} \left[e_i - \frac{\partial e_i}{\partial \phi} \Delta\phi - \frac{\partial e_i}{\partial \lambda} \Delta\lambda - \frac{\partial e_i}{\partial F} \Delta F \right] = 0 \qquad (8)$$

$$\sum_{i=1}^{m} \frac{\partial e_i}{\partial F} \left[e_i - \frac{\partial e_i}{\partial \phi} \Delta\phi - \frac{\partial e_i}{\partial \lambda} \Delta\lambda - \frac{\partial e_i}{\partial F} \Delta F \right] = 0$$

Because only linear, first-order terms are used, the values of $\Delta\phi$, $\Delta\lambda$, and ΔF which satisfy these equations will be an approximation to the exact solution. Therefore, the original estimates of latitude, longitude, and frequency are adjusted in accordance with the first solution, and new slant ranges, residuals, and partial derivatives of the residuals are computed for another solution. This process is repeated, or iterated, until the computed values of $\Delta\phi$, $\Delta\lambda$, and ΔF are sufficiently small, at which point the solution is said to have converged. Normally, only two or three iterations are required, even when the initial estimate is tens of kilometers from the final solution. Note that ignoring higher order terms has no effect on final accuracy, because these terms tend to zero as the solution converges.

In summary, the Transit position fix begins with an estimated position and determines the shift in that position required to best match calculated slant range differences with those measured by the Doppler counts. The initial estimate can be in error by 200 or 300 km and the solution will converge to an accurate value.

5.5 Accounting for Motion

If the navigator is in motion during the satellite pass, the motion must be recorded before an accurate position fix can be computed. As Figure 5 in Section 2 shows, only if the motion is known can the calculated range differences from satellite to receiver be compared properly with the range differences measured by the Doppler counts. Automatic speed and heading inputs often are employed for this purpose. During the satellite pass the computer creates a table of the navigator's estimated latitude and longitude at the beginning and end of each Doppler count interval. As before, the fix solution provides a delta-latitude and a delta-longitude, which are added to every point in the navigator's table between iterations of the solution. Therefore, although the final position fix result may be expressed as a latitude and a longitude at one point in time, the fix solution in fact is a shift of the entire estimated track.

6.0 ACCURACY CONSIDERATIONS

6.1 Static System Errors

Reference 11 presents an error budget for individual Transit position fixes that provides a good summary of the factors affecting accuracy when the navigator is not moving:

Source		Error (meters)
1.	Uncorrected propagation effects (ionospheric and trospheric effects)	1-5
2.	Instrumentation and measurement noise (local and satellite oscillator phase jitter, navigator's clock error)	3-6
3.	Uncertainties in the geopotential model used in generating the orbit	10-20
4.	Uncertainties in navigator's altitude (generally results in bias in longitude)	10
5.	Unmodeled polar motion and UT1-UTC effects	0-10
6.	Incorrectly modeled surface forces (drag and radiation pressure acting on the satellites during extrapolation interval)	10-25
7.	Ephemeris rounding error (last digit in ephemeris is rounded)	5

Since publication of this table in 1973, the polar motion error has been modeled and is included as an adjustment to the transmitted orbit parameters. The root sum square (rss) of the remaining errors lies in the range of 18 to 35 meters, which we believe is slightly optimistic due to the laboratory standards and the sophisticated refraction correction models employed by the Applied Physics Laboratory. Field results usually lie in the range of 27 to 37 meters rss. Figure 6 presented a typical set of stationary fix results. The maximum error was 77 meters, and the rss radial error was 32 meters for all 69 points.

6.1.1 Refraction Errors

There are two sources of refraction error; the larger one is due to the ionosphere. As illustrated by Figure 42, as the 150 and 400 MHz signals pass through the ionosphere, their wavelengths are stretched because of interaction with free electrons and ions. This stretching represents a phase velocity greater than the speed of light, which is characteristic of a dispersive medium. To a close first-order approximation, the wavelength stretch is inversely proportional to the square of transmitted frequency. Because satellite motion changes the path length through the ionosphere, the rate of change of this stretch causes an ionospheric refraction error frequency shift in the received signal. Reference 3 showed that an excellent refraction correction could be obtained by combining the Doppler measurements made at two different frequencies, and this is why Transit satellites transmit both 150 and 400 MHz signals.

For applications not requiring the ultimate system accuracy, 400 MHz single-channel receiving equipment can be used. Figure 42 demonstrates that because of wavelength stretching, the satellite will appear to follow a path with greater curvature about the navigator. The effect is to reduce the total Doppler shift somewhat, pushing the position fix solution away from the satellite orbit to explain the lower Doppler slope. Because the satellites move primarily along north-south lines, the resultant navigation errors are mostly in longitude. The magnitude of these errors varies with density of the ionsphere from very small at night to peaks of 200 to 500 meters in daylight, depending on sunspot activity and location with respect to the magnetic equator where the ionosphere is most dense. Figure 43 is a plot of typical single-channel results containing both daytime and nighttime fixes in which the maximum error is 242 meters and the rss error is 88 meters.

The second source of refraction error is the troposphere. In this case, propagation speed is slowed as the signal passes through the earth's atmosphere, which compresses the signal wavelength. The effect is directly proportional to transmitted frequency, as is the Doppler shift, and therefore it cannot be detected like ionospheric refraction. There are only two ways to reduce the effect of tropospheric refraction. First is by modeling its effect on the Doppler counts. Very sophisticated models employing measurements of temperature, pressure, and humidity have been published for this purpose, but less sophisticated models are usually sufficient (Reference 8). This is especially true in conjunction with the second technique, which is to delete Doppler data taken close to the horizon where the tropospheric refraction error is greatest. Above 5° to 10° of elevation, the tropospheric error is many times smaller than at the horizon, as illustrated by Figure 44 which shows typical magnitude of range error as a function of elevation above the horizon.

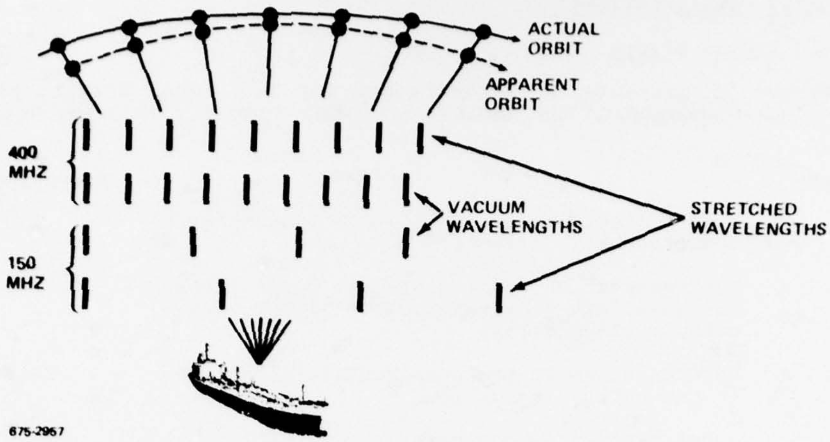


Figure 42. Ionospheric Refraction Stretches Signal Wavelength Causing Greater Apparent Orbit Curvature

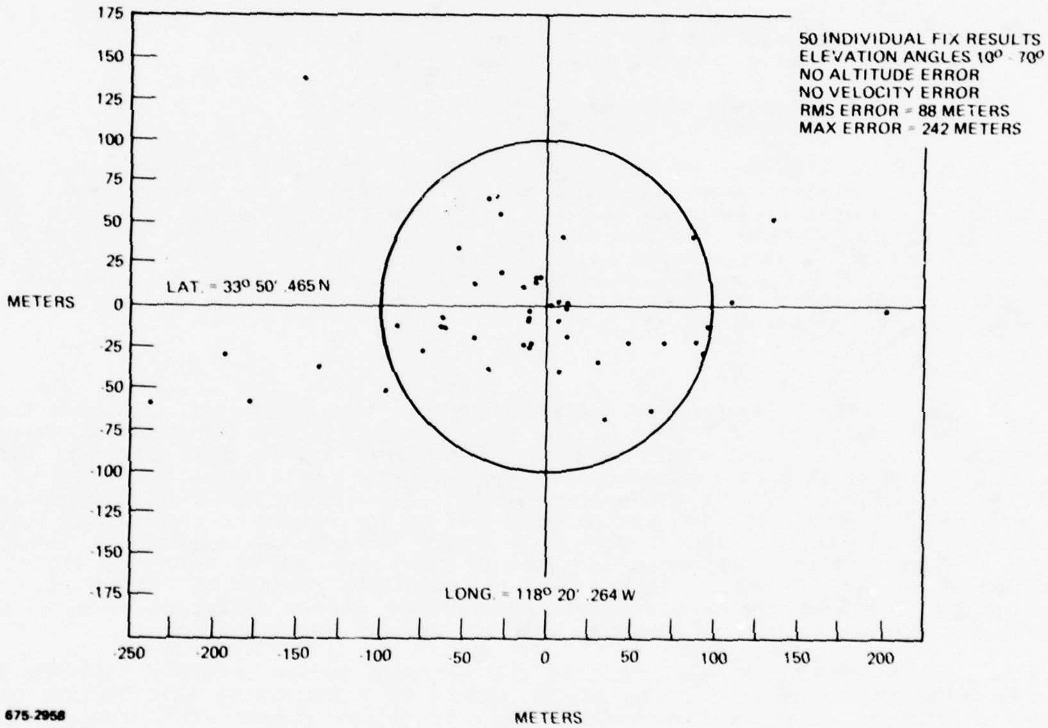


Figure 43. Typical Single-Channel Transit Position Fix Results

6.1.2 Altitude Error

The specific Doppler curve obtained as a satellite passes is predominantly a function of the navigator's position along the line of satellite motion and his distance from the orbit plane. Because Transit satellites are in polar orbits, the along-track position closely relates to latitude and the cross-track distance is a combination of longitude and altitude.

Figure 45 is the cross section of a pass where the satellite is moving in its orbital plane perpendicular to the page. It has just reached the center of pass with respect to stations X, Y, and Z. The figure illustrates how the cross-track distance is a function of both longitude and altitude, which affect the Doppler curve in similar ways. To compute an accurate fix, therefore, it is necessary to have a priori knowledge of altitude. Figure 46 shows the sensitivity of fix error to altitude error as a function of maximum satellite pass elevation angle. The elevation angle is plotted on a scale that is uniform in probability of satellite pass occurrence. In other words, more passes fall between 10° and 20° than between 70° and 80°, except at very high altitudes.

For satellite navigation "altitude" means height above or below the reference spheroid (the reference ellipsoid or satellite datum). This surface is chosen to be a worldwide best fit to mean sea level, which is the true geoid. Figure 47 illustrates the differences between the geoid, the spheroid, and topography. Therefore, knowing height above mean sea level is not sufficient for an accurate position fix. One also must know the local geoidal height, which is the deviation between the geoid and the spheroid. Figure 48 is a geoidal height map indicating that these deviations reach nearly 100 meters.

The geoidal height chart of Figure 48 was developed by observing the influence of the earth's gravity field on satellite orbits. As a result, extremely fine grain structure cannot be detected, and the map is known to be in error by ±20 meters in many places. Since the map was first published in 1967, refinements have been made but have not been released because of military classification. Thus, for maximum accuracy it is better to determine local geoidal height by the fixed site survey techniques described in Section 3.5.

6.2 Accuracy Underway

All discussion of navigation error for a stationary receiver applies equally well to a moving receiver, as long as the motion is precisely known and secondary system errors are not introduced. If the motion is not known accurately, however, additional position fix error will occur.

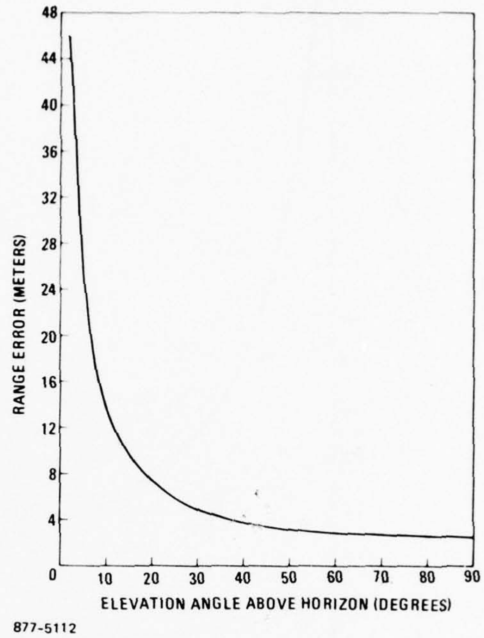


Figure 44. Typical Range Measurement Error Due to Trospheric Refraction

Figure 49 is a useful way to visualize the effect of velocity error. The ellipse is fitted through eight position fixes, with a one-knot velocity error in each of eight compass directions. Note that fix error is greater when the velocity error is in a north-south direction, and the fix error direction depends on direction of satellite motion and on whether the pass is to the east or to the west of the observer.

Whereas Figure 49 is for a single position fix, Figures 50 and 51 show the errors caused by a one-knot velocity error north and east, respectively, as a function of maximum pass elevation angle. As with Figure 46, elevation angle is plotted to represent a uniform probability of pass occurrence.

Figure 7 in Section 2 was an attempt to express overall Transit system accuracy in a single set of curves. As can be seen now, the result is overly simplified and somewhat conservative, but it does indicate realistic rss performance levels. One can see that a dual-channel system provides maximum benefit when there is an accurate source of velocity. The other benefit of the dual-channel system is to eliminate the peak 200 to 500 meter errors which occur with single-channel equipment during the day, dependent on sunspot activity.

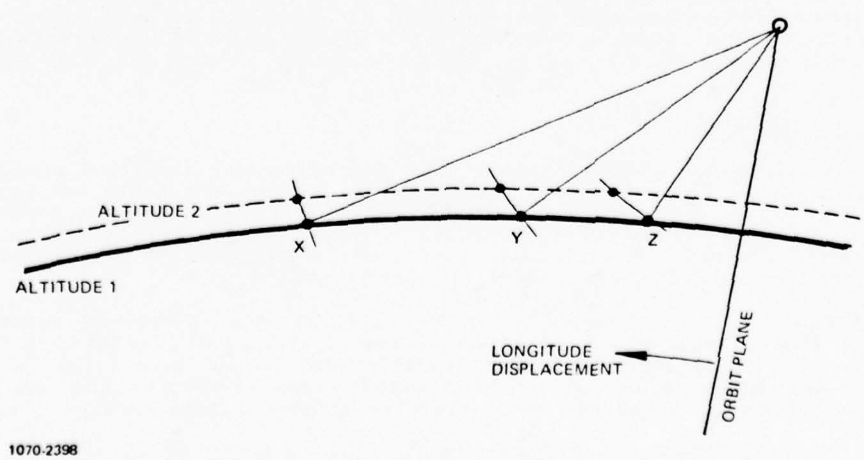


Figure 45. Effect of Altitude Estimate on Position Fix

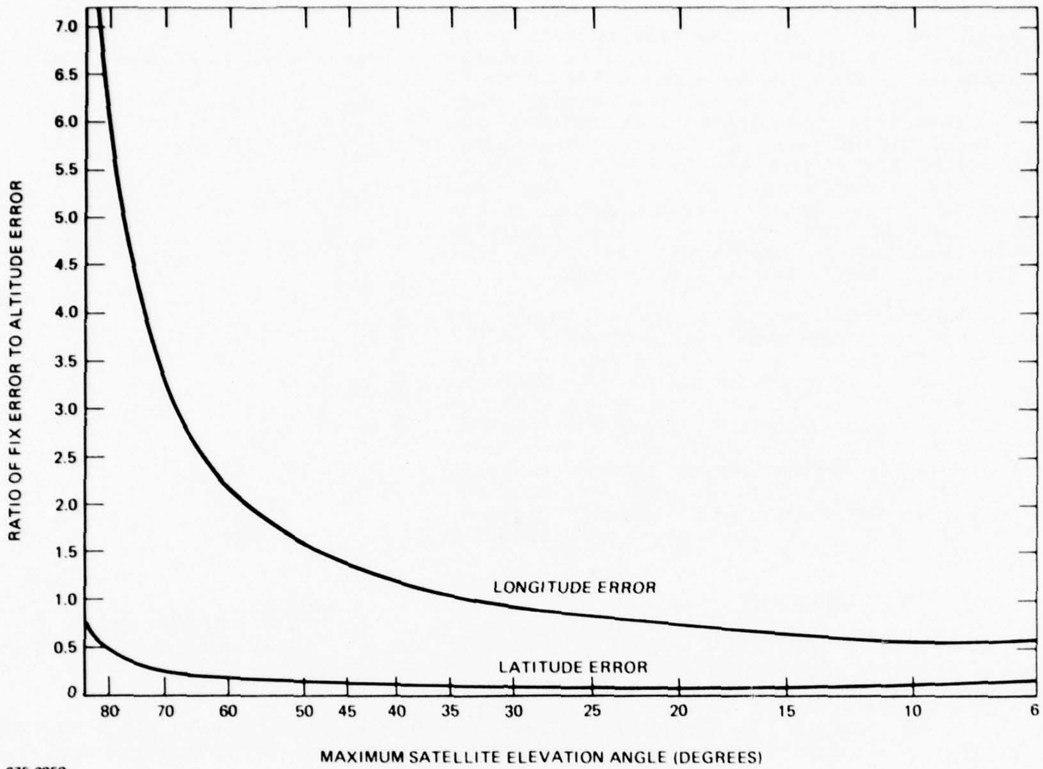


Figure 46. Sensitivity of Satellite Fix to Altitude Estimate Error

6.3 Velocity Solution

The normal position fix solution determines latitude, longitude, and frequency offset by means of Equations 8. These equations easily could be expanded to include other system variables such as velocity north, velocity east, altitude, and even acceleration. With every new variable, however, accuracy would become more and more sensitive to system noise. In fact, studies have shown that velocity north is the only parameter which can be added without creating intolerable noise sensitivity; that is, it is the only other variable which affects Doppler curve shape in a way that can be discerned clearly from the effects of latitude, longitude, or frequency. To be precise, the added variable should be velocity parallel to satellite motion, but velocity north is an adequate approximation at most latitudes because the satellites are in polar orbits.

Solving for velocity north increases position fix error when ship's motion is accurately known. Therefore it should be attempted only when velocity errors are likely to exceed about 0.4 knot. The expanded solution is more sensitive to other sources of system noise, such as asymmetric Doppler data, and it does not work well for pass elevation angles below 20°. Finally, the velocity north result becomes the scapegoat for other system errors and is not a dependable measure of velocity north error; it simply allows the latitude and longitude to be more accurate in the face of large velocity errors.

6.4 Reference Datum

It is important to realize that maps are drawn and positions are defined with respect to a reference datum. In the United States we use the North American Datum, in Japan the Tokyo Datum, in Europe the European Datum, etc. The Transit system currently uses the World Geodetic System of 1972 (WGS-72). As a result, the same reference marker will have a different set of latitude and longitude coordinates in each reference datum. Apparent differences of ½-kilometer occur in some locations.

The four parts of Figure 52 help us visualize the concept of reference datums and how they relate to each other. Figures 47 and 48 already indicated that the earth is an irregular shape due to density (gravity) variations, and Figure 52(a) is an exaggerated model of an irregular "earth". The surface shown represents the geoid, which is defined as the location of mean sea level over the entire earth's surface.

In order to make reasonably accurate maps, a model of the earth's surface is needed. Figure 52(b) shows how such models have been designed to fit the earth over the area of local interest, which in the past never was larger than a continent. The model consists of a spheroid (ellipsoid) and one position called the datum at which latitude and longitude are defined. Such a model works well and allows accurate maps to be drawn in the vicinity of the datum.

The reference oscillator output also is divided in frequency to drive the memory system. In this way, the navigation message stored there is read out and encoded by phase modulation onto both the 150 and 400 MHz signals at a constant and carefully

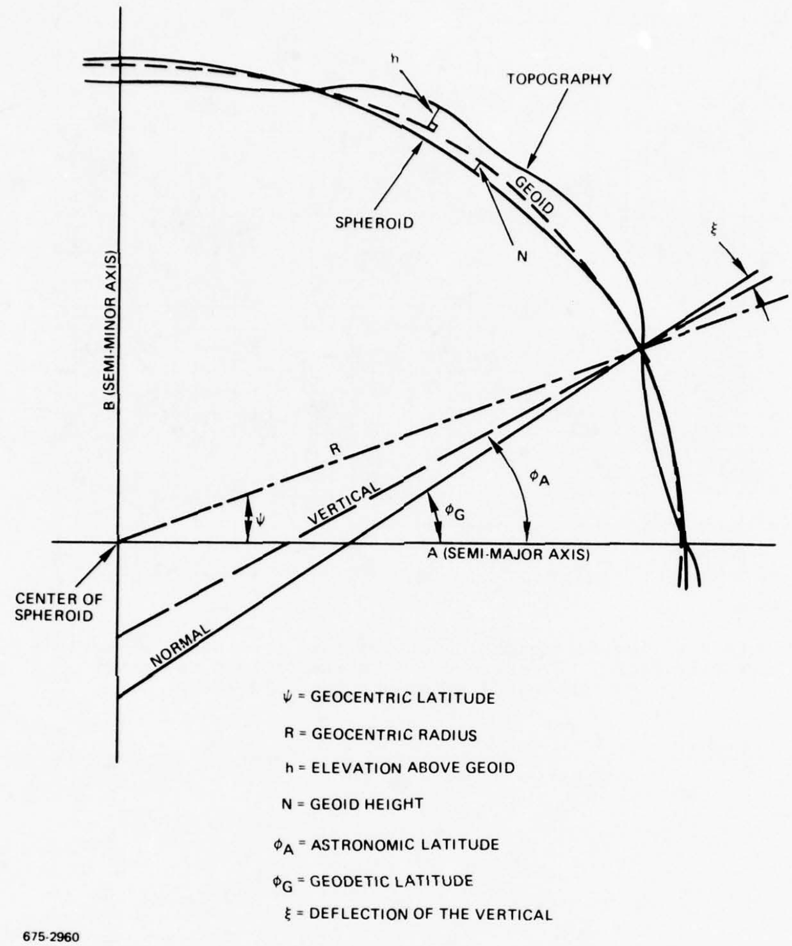


Figure 47. Relationships of Geodetic Surfaces (From NASA Directory of Observation Station Locations, 2nd Ed., Vol. 1, Nov. 1971, Goddard Space Flight Center)

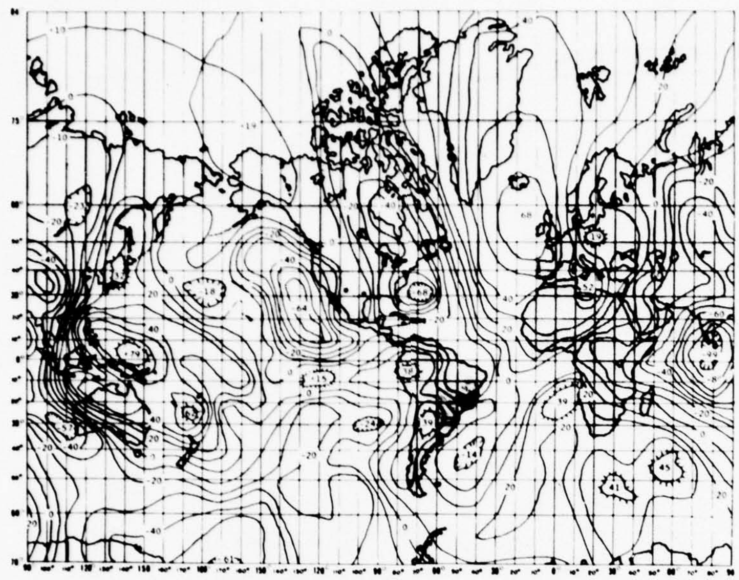
Now that satellites are being used to measure the geoid (satellite geodesy), a different type of datum is needed. As illustrated by Figure 52(c), a world spheroid may not fit the earth very well at any one location, but it is a "best fit" to the entire earth. In additon, there is not a single reference datum position because many satellite tracking stations are involved, and their positions are defined as part of the calculations which determine the earth's geopotential field (geoid). The WGS-72 spheroid is a "best fit" to the WGS-72 geoid.

Figure 52(d) makes it clear that there must be some method of relating a position in one datum to coordinates in another. For example, satellite position fixes taken in Tokyo harbor might show the ship to be well inland when plotted on a local chart. The reason is datum difference as illustrated by Figure 52(d).

The coordinate differences between two datums can be resolved by knowledge of three (or four) offset parameters and the size and shape of each spheroid. First is the x, y, and z offset between the center of the two spheroids. Sometimes a longitude rotation is needed as a fourth offset. The size and shape of each spheroid are defined by the semi-major axis (equatorial radius) and by the flattening coefficient.

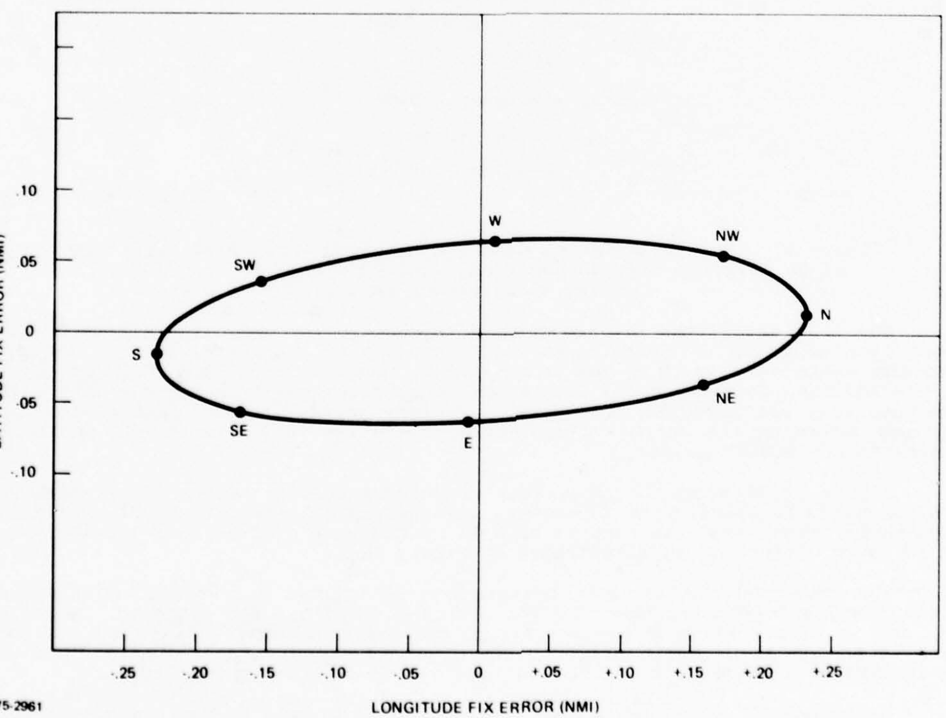
Reference 13 lists datum shift constants which can be used in converting from various datums to WGS-72, shown here in Figure 53. Caution should be exercised in trusting the results for two reasons. First is that Reference 13 indicates the accuracy of each offset constant is only ±5 meters in North America, ±10 meters in Europe, and ±15 meters in Japan and Australia. Part of this uncertainty is due to distortions in the local reference datum. The second reason is that the offset parameters were determined empirically with Geoceiver surveys using precise ephemeris orbits (see Section 3.5.4). Unfortunately, there are differences of perhaps 10 meters between positions determined with precise ephemeris orbits from the Defense Mapping Agency and those determined with

Figure 35 indicates how the navigation message defines the position of the satellite. During every two minute interval the satellite transmits a message consisting of 6,103 binary bits of data organized into 6 columns and 26 lines of 39-bit words, plus a final 19 bits. The message begins and ends at the instant of the even minute, which



375-1351

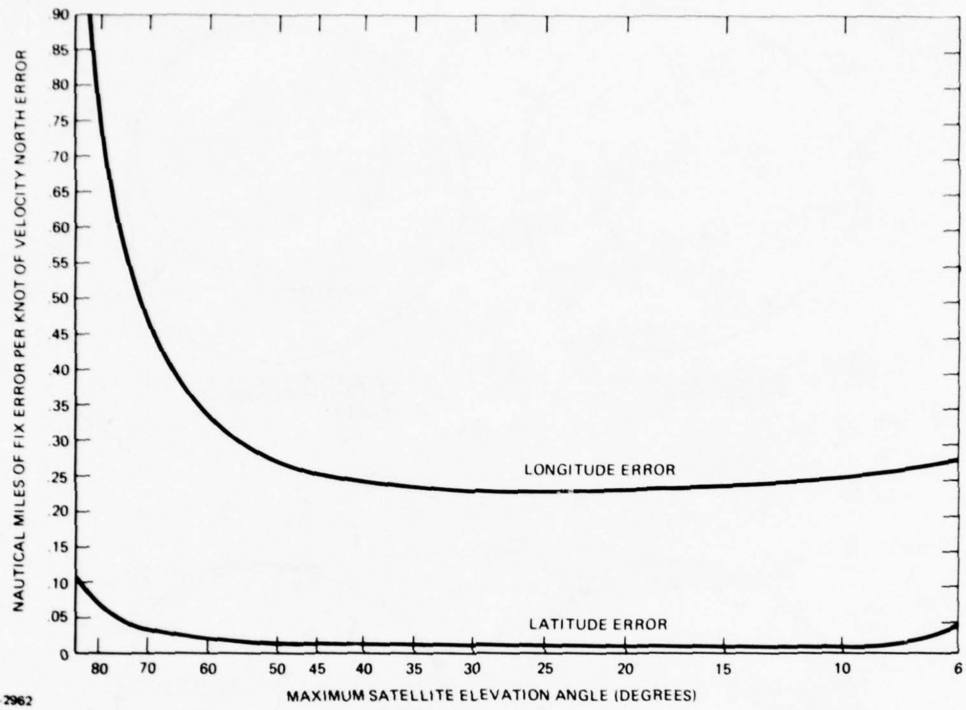
Figure 48. Geoidal Height Chart Obtained From Model of Earth's Gravity Field. Dimensions are Meters of Mean Sea Level Above the Reference Spheroid



675-2961

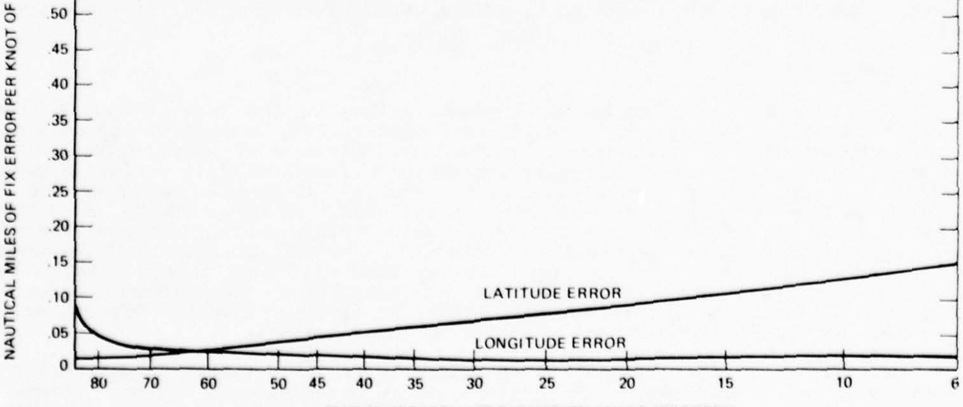
Figure 49. Effect of a One-Knot Velocity Error on the Position Fix From a 31° Satellite Pass. Direction of Velocity Error is Noted Beside Each of the 8 Fix Results. Satellite Was East of Receiver and Heading North.

past the half hour, i.e., 14 minutes or 44 minutes after the hour. This is why it is necessary to initialize a Transit set to within plus or minus 15 minutes of correct (GMT) time in order to synchronize properly. A time error of less than 15 minutes will be resolved by the "Q" number from the satellite message.



675-2962

Figure 50. Sensitivity of Satellite Fix to a One-Knot Velocity North Estimate Error



675-2963

Figure 51. Sensitivity of Satellite Fix to a One-Knot Velocity East Estimate Error

being the polar axis (mean pole of 1900-1905 or Conventional International Origin) and X being in the equatorial plane through the Greenwich meridian. The two rotations account for the longitude of the orbit plane at t_k and the inclination of the orbit with respect to earth's equatorial plane.

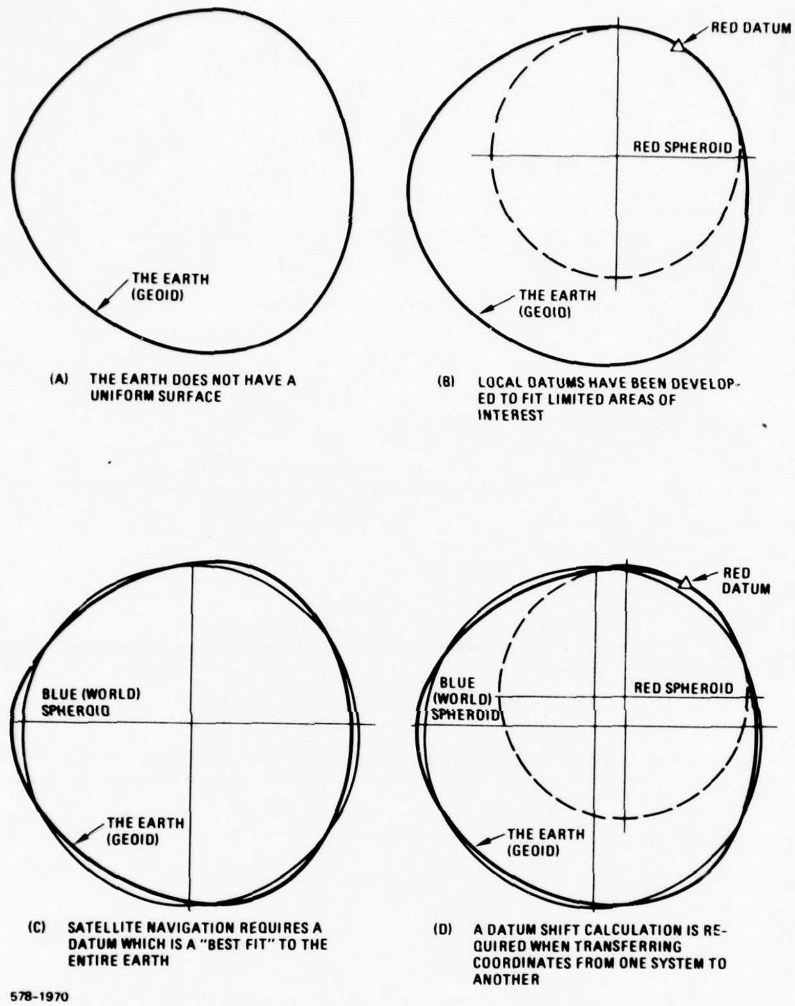


Figure 52. Development and Relationship of Local and Global Reference Datums

orbits transmitted from the Transit satellites. Figure 54, from Reference 13, gives the Molodensky formulas most often used to transform coordinates from one reference system to another.

7.0 CONCLUSION

This chapter has provided an in-depth review of the Transit system from the user's point of view. Except for a classified Soviet system, Transit is the only navigation satellite system available today. Furthermore, because of propagation limitations of the Omega system, Transit is the only system which provides truly worldwide coverage. This situation will continue until at least 1985, or later, when NAVSTAR, the Global Positioning System, is expected to become operational. As proposed by the Office of Telecommunications Policy (Reference 12), a ten-year overlap period from the time NAVSTAR becomes operational will allow users to depreciate Transit equipment before having to purchase NAVSTAR equipment. The ten-year overlap also will give time for NAVSTAR manufacturers to develop, improve, and produce a sufficient range of equipment to serve the many expected applications (Reference 23). Thus, we feel certain that Transit will continue to provide its most useful service until at least 1995.

We have shown that Transit is an extremely reliable system in delivering accurate position fixes to its users. The reliability is based on many factors. Signals are provided on a direct, line-of-sight basis from the satellite to the user, avoiding the propagation problems that plague earth-based transmitters. The Navy Astronautics Group has established a remarkable record for maintaining a reliable message in each satellite memory. The satellites themselves are extremely reliable, with three which are operating extremely well after more than ten years of service. The twelve spacecraft in storage assure that the system can be maintained in service for many years, even when the present satellites cease to function.

being received from the satellite consists of the frequency being ...
Doppler frequency shift of up to ±8 kHz due to relative motion between the satellite and the receiver. Note that the transmitted frequency is offset low by about 80 ppm (32 kHz at 400 MHz) to prevent f_R from crossing 400 MHz.

DATUM	SPHEROID	SEMI-MAJOR AXIS	RECIPROCAL FLATTENING	SHIFT TO WGS-72 a = 6378135 1/f = 298.26		
		METERS		METERS ΔX	ΔY	ΔZ
NAD 1927	CLARKE 1866	6378206	294.98	-22*	157*	176*
EUROPEAN	INTERNATIONAL	6378388	297.00	-84	-103	-127
TOKYO	BESSEL	6377397	299.15	-140	516	673
AUSTRALIAN NATIONAL	REFERENCE ELLIPSOID 1967	6378160	298.25	-122	-41	146
OLD HAWAIIAN	CLARKE 1866	6378206	294.98			
MAUI				65	-272	-197
OAHU				56	-268	-187
KAUAI				46	-271	-181
CAPE (ARC)	CLARK 1880 (MOD)	6378249	293.47	-129	-131	-282
SOUTH AMERICAN	REFERENCE ELLIPSOID 1967	6378160	298.25	-77	3	-45
ORDNANCE SURVEY OF GREAT BRITAIN 1936	AIRY	6377563	299.32	368	-120	425
JOHNSTON ISLAND ASTRO 1961	INTERNATIONAL	6378388	297.00	192	-59	-211
WAKE-ENIWETOK 1960	HOUGH	6378270	297.00			
KWAJALEIN ATOLL				112	68	-44
WAKE ISLAND				121	62	-22
ENIWETOK ATOLL				144	62	-38
WAKE ISLAND ASTRO 1952	INTERNATIONAL	6378388	297.00	283	-44	141
CANTON ISLAND ASTRO 1966	INTERNATIONAL	6378388	297.00	294	-288	-382
GUAM 1963	CLARKE 1866	6378206	294.98	-89	-235	254
ASCENSION ISLAND ASTRO 1958	INTERNATIONAL	6378388	297.00	-214	91	48
SOUTH ASIA	FISCHER 1960	6378155	298.30	21	-61	-15
NANKING 1960	INTERNATIONAL	6378388	297.00	-131	-347	0
ADINDAN	CLARKE 1880	6378249	293.47	-152	-26	212
MERCURY 1960	FISCHER 1960	6378155	298.30			
NAD 27 AREA				-25	46	-49
ED AREA				-13	-88	-5
TD AREA				18	-132	60
MODIFIED MERCURY 1968	FISCHER 1968	6378150	298.30			
NAD 27 AREA				-4	12	-7
ED AREA				-3	1	-6
TD AREA				22	34	2

*VALUES OF -9, 139, AND 173 SHOULD BE USED FOR ALASKA AND CANADA

578-1971

Figure 53. Datum Shift Constants

We have looked at the amazing breadth of Transit system applications, ranging from use aboard fishing boats to military submarines. If the user population growth trend continues, there will be more than 10,000 Transit system users by the early 1980's. Complementing the growth in applications and in the number of users is development of a new generation of Transit satellite called NOVA. Thus, there are many signs that the system is growing and fulfilling vital needs around the world.

Finally, this chapter has described both the theory of Transit satellite navigation and the factors which affect accuracy performance. This has included a definition of the orbit message parameters, the meaning of the Doppler counts, and a review of the position fix concept. The inherent system accuracy was described, and sensitivity curves were given for external factors which affect position fix accuracy.

The primary objective of this chapter has been to provide an extensive and detailed review of the Transit system today. A fascinating story has emerged. The system was developed almost exclusively to guide Polaris submarines, and it continues to serve this purpose extremely well. However, the U.S. Government also released the system for commercial use, and on their own initiative manufacturers around the world began to produce Transit navigation equipment. A wide variety of users are now experiencing the advantages of accurate, worldwide, all-weather navigation. The momentum of use continues to build, and Transit is destined to play a vital role in the world navigation scene for another decade or two.

Each Doppler count is composed of two parts: the count of a constant difference frequency $f_G - f_T$, minus the count of the number of Doppler cycles received during

A. THE STANDARD MOLODENSKY FORMULAS

$$\Delta\phi'' = \{-\Delta X \sin\phi \cos\lambda - \Delta Y \sin\phi \sin\lambda + \Delta Z \cos\phi + \Delta a\,(R_N e^2 \sin\phi \cos\phi)/a + \Delta f\,[R_M (a/b) + R_N (b/a)] \sin\phi \cos\phi\} \bullet [(R_M + H) \sin 1'']^{-1}$$

$$\Delta\lambda'' = [-\Delta X \sin\lambda + \Delta Y \cos\lambda] \bullet [(R_N + H) \cos\phi \sin 1'']^{-1}$$

$$\Delta H = \Delta X \cos\phi \cos\lambda + \Delta Y \cos\phi \sin\lambda + \Delta Z \sin\phi - \Delta a\,(a/R_N) + \Delta f\,(b/a)\, R_N \sin^2\phi$$

B. THE ABRIDGED MOLODENSKY FORMULAS

$$\Delta\phi'' = [-\Delta X \sin\phi \cos\lambda - \Delta Y \sin\phi \sin\lambda + \Delta Z \cos\phi + (a\Delta f + f\Delta a) \sin 2\phi] \bullet [R_M \sin 1'']^{-1}$$

$$\Delta\lambda'' = [-\Delta X \sin\lambda + \Delta Y \cos\lambda] \bullet [R_N \cos\phi \sin 1'']^{-1}$$

$$\Delta H = \Delta X \cos\phi \cos\lambda + \Delta Y \cos\phi \sin\lambda + \Delta Z \sin\phi + (a\Delta f + f\Delta a) \sin^2\phi - \Delta a$$

C. DEFINITION OF TERMS IN THE MOLODENSKY FORMULAS

ϕ, λ, H = GEODETIC COORDINATES (OLD ELLIPSOID)

ϕ = GEODETIC LATITUDE. THE ANGLE BETWEEN THE EARTH'S EQUATORIAL PLANE AND THE ELLIPSOIDAL NORMAL AT A POINT (MEASURED POSITIVE NORTH FROM THE EQUATOR, NEGATIVE SOUTH).

λ = GEODETIC LONGITUDE. THE ANGLE BETWEEN THE PLANE OF THE GREENWICH MERIDIAN AND THE PLANE OF THE GEODETIC MERIDIAN OF THE POINT (MEASURED IN THE PLANE OF THE EQUATOR, POSITIVE EAST FROM GREENWICH).

H = THE DISTANCE OF A POINT FROM THE ELLIPSOID MEASURED ALONG THE ELLIPSOIDAL NORMAL THROUGH THE POINT.

$H = N + {}^*h$

N = GEOID-ELLIPSOID SEPARATION. THE DISTANCE OF THE GEOID ABOVE (+N) OR BELOW (-N) THE ELLIPSOID.

$^*h$ = DISTANCE OF A POINT FROM THE GEOID (ELEVATION ABOVE OR BELOW MEAN SEA LEVEL).

$\Delta\phi, \Delta\lambda, \Delta H$ = CORRECTIONS TO TRANSFORM THE GEODETIC COORDINATES FROM THE OLD DATUM TO WGS.

$\Delta X, \Delta Y, \Delta Z$ = SHIFTS BETWEEN ELLIPSOID CENTERS OF THE OLD DATUM AND WGS.

a = SEMIMAJOR AXIS OF THE OLD ELLIPSOID.

$^*b$ = SEMIMINOR AXIS OF THE OLD ELLIPSOID.

$^*b/a = 1 - f$

f = FLATTENING OF THE OLD ELLIPSOID.

$\Delta a, \Delta f$ = DIFFERENCES BETWEEN THE PARAMETERS OF THE OLD ELLIPSOID AND THE WGS ELLIPSOID (WGS MINUS OLD).

e = ECCENTRICITY.

$e^2 = 2f - f^2$

R_N = RADIUS OF CURVATURE IN THE PRIME VERTICAL

$R_N = a/(1 - e^2 \sin^2\phi)^{1/2}$

R_M = RADIUS OF CURVATURE IN THE MERIDIAN.

$R_M = a(1 - e^2)/(1 - e^2 \sin^2\phi)^{3/2}$

NOTE: ALL Δ-QUANTITIES ARE FORMED BY SUBTRACTING OLD ELLIPSOID VALUES FROM WGS ELLIPSOID VALUES.

*INDICATES PARAMETERS WHICH DO NOT APPEAR IN THE ABRIDGED MOLODENSKY FORMULAS.

578-1972

Figure 54. Datum Shift Equations (From References 8 and 13)

SELECTED REFERENCES

1. Black, H.D. et al, "Planned Improvements in the Transit System (1975)", Navigation, Vol. 22, No. 4, Winter 1975-76.

2. Etherington, M. "The Use of the Transit System in Tuna Fishing", Marine Technology Symposium, United States Trade Center, Mexico City, March 1977.

3. Guier, W.H. and Weiffenbach, G.C. "A Satellite Doppler Navigation System", Proc. IRE, Vol. 48, No. 4, April 1960.

4. Hatch, R.R. "Point Positioning and Translocation via the Transit Satellite System", Magnavox Report MX-TM-3220-76, presented at The 46th Annual Meeting of the Society of Exploration Geophysicists, Houston, Texas, October 1976.

5. Hatch, R.R. "New Positioning Software from Magnavox", Magnavox Report MX-TM-3217-76, presented at the International Geodetic Symposium on Satellite Doppler Positioning, Las Cruces, New Mexico, October 1976.

6. Laurila, S.H. Electronic Surveying and Navigation, Chapter 28, John Wiley & Sons, 1976, ISBN 0-471-51865-4.

7. Leroy, C.F. "Results from Portable Doppler Receivers Using Broadcast and Precise Ephemerides", Proceedings International Geodetic Symposium, Vol. I, cosponsored by U.S. Defense Mapping Agency, National Ocean Survey, NOAA, Las Cruces, N.M., October 1976.

8. Moffett, J.B. "Program Requirements for Two-Minute Integrated Doppler Satellite Navigation Solution", Johns Hopkins University Applied Physics Laboratory Report TG 819-1 (Rev. 2), 1973.

9. Moyer, J.R. "Navigation and Positioning of Drill Rigs and Other Offshore Platforms", Magnavox Report MX-TM-3248-77, Presented at Offshore Exploration Technical Symposium, Gdansk, Poland, June 1977.

10. Moyer, J.R. "The Satellite Doppler Survey System — A Modern Efficient Tool for the Field Surveyor", Magnavox Report MX-TM-3247-77, Presented at the Seminar on the Doppler System and its Application in the Determination of Stations of Geodetic Control, Buenos Aires, Argentina, July 1977.

11. Piscane, V.L. et al "Recent (1973) Improvements in the Navy Navigation Satellite System", Navigation, Vol. 20, No. 3, Fall 1973.

12. "Radio Navigation Systems Economic and Planning Analysis, Final Report", Volumes 1, 2, and 3, Prepared for Office of Telecommunications Policy by Computer Sciences Corporation, July 1977.

13. Seppelin. T.O. "The Department of Defense World Geodetic System 1972", Defense Mapping Agency, Washington, D.C., May 1974.

14. Sharpe, R.T. and Galyean, P.G. "An Acoustic Navigation System for Site Survey and Geodetic Positioning", Magnavox Report MX-TM-3236-77, Presented at the Offshore Technology Conference, Houston, Texas, May 1977.

15. Sharpe, R.T. "An Integrated Acoustic Positioning System", Navigation, Vol. 24, No. 3, Fall 1977.

16. Stansell, T.A. "The Navy Navigation Satellite System: Description and Status", Navigation, Vol. 15, No. 3, Fall 1968.

17. Stansell, T.A. "Transit, The Navy Navigation Satellite System", Navigation, Vol. 18, No. 1, Spring 1971.

18. Stansell, T.A. "Accuracy of Geophysical Offshore Navigation Systems", Offshore Technology Conference Preprints, OTC 1789, Dallas, Texas, April-May 1973.

19. Stansell, T.A. "Achieving Reliability in Automatic Navigation Equipment," Second International Symposium on Ship Operation Automation, Washington, D.C., August 1976.

20. Stansell, T. A. "Doppler Survey Equipment: Background, Requirements and Trends", Magnavox Report MX-TM-3219-76, Presented at the International Geodetic Symposium on Satellite Doppler Positioning, Las Cruces, New Mexico, October 1976.

21. Stansell, T.A. "Positioning and Navigation by Satellite", Magnavox Report MX-TM-3261-77, Joint Conference on Satellite Applications to Marine Operations, New Orleans, Louisiana, 15-17 November 1977.

22. Stansell, T.A. "The Many Faces of Transit", Navigation, Vol. 25, No. 1, Spring 1978.

23. Stansell, T.A. "Civil Marine Applications of the Global Positioning System", Navigation, Vol. 25, No. 2, Summer 1978.

The Timation Navigation Satellites

Roger L. Easton, Code 7960
Head, Space Applications Branch
U. S. Naval Research Laboratory
4555 Overlook Ave., S. W.
Washington, D. C. 20375

SUMMARY

The development of stable RF oscillators in the two post World War II decades was essential to the design of passive ranging navigation systems. Early involvement of the U. S. Naval Research Laboratory in this method of time synchronized navigation is outlined. The first NRL satellite used to validate the concept was launched in 1967. Both it and the second satellite used quartz crystal oscillators. The two following satellites use rubidium and cesium standards respectively.

The lessons learned from the early satellites are summarized, and the merger of TIMATION into the NAVSTAR Global Positioning System Program is described. Selected experimental results from NTS-1 and 2 are analyzed and the plans for NTS-3 are outlined.

1.0 INTRODUCTION

1.1 Early Background

The technology base for the original Timation experiments was developed at the Naval Research Laboratory during the late 1950s in the search for a practical method of tracking the early U. S. space satellites. The Minitrack network as it was called, used RF interferometry requiring highly stable frequencies as its tracking parameter and this technique led to the idea of position fixing using synchronized frequency standards to measure the range from satellite to observer. Distance measurement by time synchronization was further explored during the development of the Navy's Space Surveillance System (WS-434) in 1958-64 using early cesium standards. Thus the anagram TIMe navigATION was born and the concept took shape, as the atomic clock technology necessary to support such a navigation system developed during the 1960s and early 70s. The navigation satellite clock development program is described in this NATO AGARDOGRAPH by C. A. Bartholomew in his article entitled "Evolution of Frequency Standards". The NRL TIMATION project began with a task from the Bureau of Naval Weapons in September 1964. This work is currently sponsored by the Naval Electronics Systems Command (PME-106).

1.2 The TIMATION Navigation Concept

The basic method of position fixing by means of man-made satellites is as old as the science of celestial navigation except for one minor particular. In the case of star navigation, the distances are not known but the angles may be calculated in accordance with the laws of Newton and Kepler so that they can be predicted in advance and reduced to tabular form for use in calculating lines-of-position. In the case of satellite navigation, distances are known and are used as the basis for solving the navigation equation. On figure 1 we see the basic navigation triangle. The height of the satellite above the earth's center is determined within a few meters. The earth's radius is also known and the range from the observer to the satellite is measured electronically. Since the other sides of the triangle are known, the range line serves to describe a line-of-position on the surface of the earth upon which the observer must be located. Two such lines-of-position determine a two dimensional fix. Three are needed to determine a position plus an altitude in the case of an aircraft. Additional lines of position, if more than three satellites are available, can be used to provide more accuracy or to measure time.

The next two figures show the method of determining a fix based on lines-of-position from satellites. With modern computer technology, this can all be done automatically so that the aircraft pilot or ship's captain can read his position continuously from a dial. The second graph shows the effect of a timing error which results in the equivalent of a "large triangle" in conventional navigation terminology.

Figure 4 shows the basic Timation geometry as seen from a ship. Precise ranges are measured from two or more satellites to determine a fix. To make use of this form of navigation, accurate orbits must be known for the satellite, precise clocks must be available and the clocks in the satellites must be synchronized with the clock of the observer.

2.0 TIMATION I

2.1 Description

The first satellite experiment to demonstrate the validity of this passive ranging technique was launched into a 500 mile high orbit on May 31, 1967. It weighed 85 lbs,

produced 6 watts of electrical power, and contained a quartz clock capable of maintaining its stability within 3 parts 10^{11} ($\Delta f/f$) per day. The satellite, ready for launch, is shown in figure 5. Figure 6 shows an artist's conception of the satellite with the gravity gradient boom and solar panels extended.

2.2 Timation I Experiments

During June 1967, this TIMATION I satellite demonstrated that lines of position could be determined from ranging satellites and fixes from a single satellite could be obtained using range and doppler measurements. Figure 7 shows a Timation I fix for a ship taken on May 15, 1968. The fix and the charted position disagreed by 0.2 NM. Figure 8 shows a Timation I position of an aircraft which differed by 0.3 NM from the actual track. Timation I was also used to transfer time between the Laboratory and Ft. Collins, Colorado; Crane, Indiana; and Sanford, Florida. The accuracy of these time transfers was consistantly below 1 microsecond.

Timation I results displayed significant errors from two sources; solar radiation and ionospheric refraction. Initially it was suspected that electrons produced most of the errors produced by the ambient radiation environment. For this reason, Timation II's crystal oscillators were surrounded by 1/8" lead shield.

3.0 TIMATION II

3.1 Description

The second Timation was launched into a 500 mile orbit on September 30, 1969. It also contained a quartz oscillator, however the stability was better, about 1 part in 10^{11}. It weighed 125 lbs and produced three times the power of Timation I. Figure 9 shows the TIMATION II Thor Agena resting on its pad ready for launch. This satellite was also launched on a vehicle carrying a high priority military payload. The external appearance of the satellite was similar to that of TIMATION I.

3.2 Proton/Ionospheric Effects

A large frequency shift in the Timation II clocks was observed during the early part of August 1972 and this coincided with a solar proton storm. From this happy accident it was determined that protons instead of electrons were the major cause of quartz clock frequency shifts. This phenomenon is shown in figure 10.

Ionospheric effects were investigated with Timation II by using dual frequency coherent transmissions from the satellite. This allowed measurement of the ionospheric delay and, of considerable scientific interest, displayed its fine structure as the satellite passed over the observer.

Figure 11 shows a block diagram of the equipment used to determine the ionospheric scintillations with TIMATION II and figure 12 shows a comparison of delays derived from measurements of the electron content and those achieved from TIMATION II range measurements. From these and other data it was determined that the ionosphere can have appreciable irregularities and can vary tremendously in total ionospheric content over a few days time.

3.3 Navigation Experiments

Several techniques were used in making navigation fixes. One can obtain several range lines of position, advance them to the same time and thereby obtain a fix. One can do the same with doppler or use a combination of a range and doppler to obtain nearly instantaneous fixes from a single satellite. The major problem with doppler arises from the fact that it is sensitive to the user's velocity. Only for a user who knows his velocity accurately (or keeps it constant for several measurements) is the doppler measurement method a practical one.

Other parameters that were measured with Timation II were position accuracy as a function of maximum modulation frequency and ionospheric correction.

Figure 13 shows the accuracy of fixes determined by Timation II as a function of maximum modulation frequency read out at two stations. It shows that there is but a slight difference in the results from a quiet station (Ft. Valley) and a noisy station (Chesapeake Bay or CBD). This graph also illustrates that the resolution level was still limiting the system accuracy.

Another comparison was also made - that of a single (400 MHz) frequency versus a dual frequency observation. Figure 14 shows this comparison for three cases. The best accuracy was obtained with the dual frequency in which the effect of the ionosphere was measured and removed. Second best was the modeled ionosphere and poorest was the single frequency which shows an error of approximately 1 microsecond. For the single frequency case it was apparent that there is no accuracy improvement as one goes from the 300 KHz to the 1 MHz modulation. The reason is that for this case the accuracy was limited by the ionosphere - not by the resolution.

In addition to measuring range lines of position it was possible to measure range rate or doppler lines of position with Timation II. Figure 15 shows both range and

range-range rate fixes. The range-range rate fix is made from a single range LOP and a single range rate LOP while the range only fix uses a complete satellite pass during which 10 LOPs were obtained. As one might suspect, the range only fixes demonstrated better accuracy than the range-range rate fixes.

4.0 NAVIGATION TECHNOLOGY SATELLITE I

4.1 Description

The final satellite in the Timation series was launched on the 14th of July 1974, after the program had been merged with the Air Force 621B project to form the NAVSTAR Global Positioning System (GPS). This spacecraft was renamed Navigation Technology Satellite One (NTS-1) and became the first of a new series of satellites launched by NRL to provide technical support for the GPS Joint Service Project Office. Figure 16 shows the configuration of NTS-1. The executive service for the GPS program is the U. S. Air Force.

4.2 NTS-1 Experiments

Numerous changes were introduced in Timation III/NTS-1 as a result of lessons learned from the earlier Timations. An altitude of 7,500 NM was selected to reduce errors due to atmospheric drag and more nearly reflect the environment of operational satellites which would be at even higher altitudes. The weight was increased to 650 lbs and the power to 125 watts. In addition to the UHF frequencies of the earlier satellites, an L-band Pseudo Random Noise (PRN) signal at the same frequency as that of the operational GPS satellites was added. Perhaps the most significant change was the addition of two rubidium clocks. These had become available about eight months prior to the launch date and were readied for operation in space by an intensive effort. The rubidiums performed well in NTS-1, demonstrating a stability of about one part in 10^{12} per day. As a result, they are to be used as the primary standards in the early Navigation Development Satellites (NDS) built by Rockwell International.

In addition to its clocks NTS-1 contained a number of other experiments, all tailored to meet requirements of the GPS. It space qualified solar cells which should increase the power and reliability of future navigation satellites.

5.0 THE NTS-2 SATELLITE

NTS-2 was the first satellite completely designed and built under the sponsorship of the NAVSTAR GPS program. The altitude selected was 10,980 NM, chosen because it coincided with the semi-synchronous heighth planned for the operational GPS satellites. Instead of rubidium, NTS-2 contained two cesium clocks, designed and built by Frequency and Time Systems, Inc. and the frequency stability consequently increased to 1-2 parts in 10^{13} per day. In addition, NTS-2 contains a duplicate of the navigation system planned for use in the NDS satellites. Both the navigation system and the clocks have performed excellently in NTS-2, and data is now being accumulated for use in the operational satellites and to checkout the ground stations. Figure 17 shows NTS-2 and the major objectives planned for this experiment.

6.0 THE NTS-3 SATELLITE

Looking into the future, an NTS-3 is planned for launch in 1981. This satellite will contain both hydrogen maser and an advanced cesium clock. The clock stability is expected to be an order of magnitude better than that of NTS-2, and the power will be increased to 450 watts. NTS-3 will also fly other advanced technology experiments to qualify new components for later operational use.

PRINCIPLE OF OPERATION OF NAVSTAR AND SYSTEM CHARACTERISTICS

Robert J. Milliken and Curt J. Zoller*
Rockwell International Corporation
Space Division
12214 Lakewood Boulevard
Downey, California, USA 90241

SUMMARY

The NAVSTAR/Global Positioning System (GPS) will provide extremely accurate three-dimensional position and velocity information to users anywhere in the world. The position determinations are based on the measurement of the transit time of RF signals from four satellites of a total constellation of 24. Accuracies on the order of 10 meters may be anticipated. This paper discusses the basic technique by which the system operates, the navigation signal, the measurement of the transit time, error sources, accuracies, and other characteristics of the system.

INTRODUCTION

The baseline constellation of 24 satellites operates in 12-hour orbits at an altitude of 20,183 km (10,898 nmi). It will provide visibility of 6 to 11 satellites at 5 degrees or more above the horizon to users located anywhere in the world at any time. Signals are transmitted at two L-band frequencies (1227 and 1575 MHz) to permit corrections to be made for ionospheric delays in signal propagation time. The signals are modulated with two codes: P, which provides for precision measurement of time, and C/A, which provides for easy lock-on to the desired signal. The satellites employ a shaped-beam antenna that radiates near-uniform power to system users of at least -163 dBW for the L_1 P-code and -160 dBW for the L_1 C/A code. The corresponding L_2 power level carrying only the P-code is at least -166 dBW. Navigation fixes can be made in a time interval of from tens of seconds to several minutes, depending on the sophistication of the receiving system.

NAVIGATION TECHNIQUE

Four satellites are normally required for navigation purposes, and the four offering the best geometry can be selected manually or automatically by receivers using ephemeris information transmitted by the satellites. Ranges to the four satellites are determined by scaling the signal transit time by the speed of light. The transmitted message contains ephemeris parameters that enable the user to calculate the position of each satellite at the time of transmission of the signal.

Operation of the system requires precise synchronization of space vehicle (SV) clocks with "GPS system time," which is accomplished by the use of an atomic frequency standard in each space vehicle and use of clock correction parameters that are provided by the Control Segment. The requirement for users to be equipped with precision clocks is eliminated by the use of range measurements from four satellites. If users maintained precision clocks synchronized with GPS system time, navigation could be accomplished with only three satellites. In that case, the user could be thought of as being at the intersection of three spheres, with centers located at the satellites. The fourth satellite permits an estimate of the user's clock error. In this case, the user position equations contain four unknowns consisting of position in three dimensions and the error, or fixed bias, in the user's imprecise clock, which can be solved by simultaneous solution of the four equations. A discussion of the solution follows. A mathematical derivation is included in the last section of the paper.

The measurement of range to the satellites, made by the user with an imprecise clock, is called "Pseudo-range" because it contains a bias of fixed magnitude in each range estimate due to the clock error.

The pseudo-range is defined as

$$\bar{R}_i = R_i + C\Delta t_{Ai} + C\,(\Delta t_u - \Delta t_{Si})$$

where

$\bar{R}_i$ = pseudo-range to the satellite

R_i = true range

C = the speed of light

Δt_{Si} = satellite i clock offset from GPS s stem time

Δt_u = user clock offset from GPS system time

Δt_{Ai} = propagation delays and other errors

Pseudo-range is illustrated in Figure 1.

* Appreciation is expressed to Mr. Edward M. Lassiter, of Aerospace Corporation, who provided source material for this paper, and to Mr. Paul S. Jorgensen, also of Aerospace Corporation, from whose work some of the material has been drawn.

The GPS user measures the apparent (pseudo-range) transit time by measuring the phase shift of identical pseudo-random noise (PRN) codes that are generated in both the space vehicle and the user receiver, each synchronized with its own clock. The receiver code is shifted until maximum correlation is achieved between the two codes; the time magnitude of the shift is the receiver's measure of pseudo-range time.

The concept of finding a user position based on the removal of the fixed range bias from each range estimate is illustrated in a two-dimensional situation with three satellites in Figure 2. The figure illustrates that the pseudo-range radii from the three satellites do not meet at a point but enclose the shaded triangular area. However, a range value of fixed magnitude ($\Delta t_u C$) can always be found that when removed from the pseudo-ranges (or added, as the case may be) will cause the radii to meet at a point, which is the user position. The value of $\Delta t_u C$ represents the range equivalent of the user clock error. If other errors (of unequal magnitude) exist in the pseudo ranges, which we will call independent errors, a value of fixed magnitude can still always be found that when removed from the pseudo-ranges will cause the radii to meet at a point. In this case, the point is an estimate of the user position that differs from the true user position by an error that is a function of the independent range errors. Similarly, the fixed magnitude that is removed then provides an estimate of the equivalent user clock error, which differs from the true user clock error by an amount that is also a function of the independent errors. The illustration in Figure 2 represents only a two-dimensional situation in which the three satellites and user all lie in the same plane. However, the same statements and logic apply in a three-dimensional situation when ranges to four satellites are used.

GPS TIME STANDARD

In terms of navigation accuracy, one nanosecond of time error is equivalent to approximately 0.3 meters (0.984 ft) of range error so that precision timing and frequency control are essential to the GPS system. All system timing requirements are synchronized with GPS system time, which is maintained by the Master Control Station (MCS) through the use of a set of highly accurate cesium clocks. Precision timing is maintained in the space vehicles by the use of a highly stable atomic clock in each vehicle with a known or predictable offset from GPS system time. The MCS monitors the SV time standards daily with reference to GPS system time and generates clock correction parameters for transmission to the space vehicles where they are retransmitted to users with the navigation signals and used to determine the precise magnitude of the clock offsets.

GPS system time necessarily differs from UTC (Universal Coordinated Time), which must be adjusted for leap seconds at periodic end-of-year intervals. Such adjustments in GPS time would disrupt the continuous availability of the space vehicles for navigation purposes. Knowledge of the difference between GPS system time and UTC is maintained within 100 microseconds, however, and the difference is published regularly for the benefit of users interested in the use of GPS as a time standard. Time users with known positions and highly stable frequency standards can determine GPS time by the use of signals from only one space vehicle. Accuracies can be on the order of 15 to 20 nanoseconds.

The space vehicle clock frequency is nominally 10.23 MHz, which is offset slightly to a center frequency of 10.22999999545 MHz to allow for relativity effects. Its maximum allowable uncertainty is one part in 10^{12} per day. The MCS has the capability of adjusting both the clock time phase and frequency, if required. The phase can be adjusted to a resoltuion of one chip ($\approx$98 nanoseconds). The frequency can be set in steps no smaller than 4×10^{-12} delta f/f over a range of ± 2 parts in 10^9 around the center frequency. Upload of clock correction parameters into the space vehicles and adjustment of the clocks is accomplished by uplink commands. The pseudo-random noise codes, which are synchronized with space vehicle time, are maintained within 976 microseconds of GPS system time in order to preclude secondary control problems such as almanac word-length limitation that would otherwise arise.

All frequencies in the space vehicle are derived from, and synchronized with, integrals of the basic 10.23 MHz SV frequency standard. These include.

		Repeat Interval or Frequency
P-code:	Reset	7 days
	Frequency	10.23 MHz
C/A code:	Epoch	1 millisecond
	Frequency	1.023 MHz
X1 Epoch (Z-count change)		1.5 seconds
HOW (handover-word) change		6 seconds
Data bit stream frequency		50 bps
L_1 RF frequency		154 x 10.23 = 1575.42 MHz
L_2 RF frequency		120 x 10.23 = 1227.6 MHz

SIGNAL STRUCTURE AND THE MEASUREMENT OF TIME

The navigation signal transmitted from the space vehicles consists of two RF frequencies, L_1 at 1575.42 MHz and L_2 at 1227.6 MHz. The L_1 signal is modulated with both the P and the C/A pseudo-random noise codes in phase quadrature. The L_2 signal is modulated with the P-code. Both the L_1 and L_2 signals are also continuously modulated with the navigation data-bit stream at 50 bps. The functions of the codes are twofold: (1) identification of space vehicles, as the code patterns are unique to each space vehicle and are matched with like codes generated in the user receiver, and (2) the measurement of the navigation signal transit time by measuring the phase shift required to match the codes. The P-code is a long precision code operating at 10.23 Mbps but difficult to acquire. The C/A (clear access) code is a short code, readily acquired, but operating at 1.023 Mbps, which provides a grosser measurement of time. The C/A code is normally acquired first and a transfer is made to the P-code by the use of the handover word (HOW) contained in the navigation data stream (see the Navigation Message). It is possible, however, for users with precision clocks precisely synchronized with GPS time and the approximate knowledge of their position (10,000 ft - 20,000 ft) to bypass the C/A code and acquire the P-code directly.

The P-code generated in each space vehicle is a pseudo-random noise chip sequence of seven days in length.(1) That is, the pattern repeats only once very seven days. The code is initialized at midnight each Saturday. GPS system time is counted from the initialization of the P-code each week. Counting is accomplished by a count of the epochs (recurrences of the initial state) of a subsidiary code generator designated X1, used in the generation of the P-code, that occur every 1.5 seconds. The count of the X1 epochs, termed Z, rises to 403,199 at the end of each week, when it is reinitialized at zero. The system time of the week is transmitted to users every six seconds in the form of the handover word, HOW.

In order for the ground receiver to lock onto the P-code, it must know approximately what time-slice in the seven-day code to search. At typical receiver search rates, on the order of 50 bits per second, the time required to search as much as one second of the seven-day P-code would require many hours. It is therefore necessary to resort to the C/A code for initial code match and lock-on.

The C/A code is a pseudo-random noise chip stream unique in pattern to each space vehicle that repeats every millisecond. It is relatively easy for the receiver to match and lock onto the C/A code because the search is limited to the time interval of one millisecond and the chip rate is only one-tenth that of the P-code. After lock-on to the C/A code, the transfer to the P-code is facilitated by the HOW word. Its change, which is synchronized with the P-code, indicates the point in the incoming P-code that will occur at the next change, i.e., within the next six seconds. The receiver-generated P-code is shifted in phase to synchronize with the designated point in the incoming P-code when triggered by the change in the HOW. The total phase-shift required for lock-on is the measured pseudo-range time, including the offset in the user clock as well as the propagation delays and system errors. The P-code frequency affords the degree of accuracy required for the measurement of signal transit time that the C/A code frequency could not. Use of the P-code also avoids ambiguity in the C/A code epoch, which repeats every millisecond, if the user clock offset exceeds this amount.

THE NAVIGATION MESSAGE

The navigation message contains the data that the user's receiver requires to perform the operations and computations for successful navigation with the GPS. The data include information on the status of the space vehicle: the time synchronization information for the transfer from the C/A to the P-code; and the parameters for computing the clock correction, the ephemeris of the space vehicle and the corrections for delays in the propagation of the signal through the atmosphere. In addition, it contains almanac information that defines the approximate ephemerides and status of all the other space vehicles, which is required for use in signal acquisitions. The data format also includes provisions for special messages.

The navigation message is formatted in five subframes of six seconds in length, which make up a data frame of 30 seconds, 1500 bits long. The data are nonreturn-to-zero (NRZ) at 50 bps and are common to the P- and C/A signals on both the L_1 and L_2 channels. The format of the data is shown in Figure 3.

Each data subframe starts with a telemetry word (TLM) and the C/A to P-code handover word (HOW). The latter permits the C/A to P transfer to be made at the termination of any six-second subframe. The initial 8 bits of the TLM contain a preamble that facilitates acquiring the data message. The balance of the TLM contains information designed primarily for the use of the Control Segment in determining the accuracy with which the daily update of the space vehicles has been received and utilized and is ordinarily not decoded by the user receivers. A space vehicle health-status word is provided in the fifth subframe (Data Block III). which indicates the status of the SV and permits the user the option of selecting another SV. The HOW is the second word of each subframe and occupies 30 bits, including parity. It starts with the uppermost significant 17 bits of the 19-bit Z-count word, which is the running indicator of time in the space vehicle. The Z-count changes every 1.5 seconds; the HOW every 6 seconds. Bit number 18 is used to indicate when a roll-momentum dump has occurred since the last upload. Momentum dumps have some affect on the ephemeris accuracy and this information provides the option for a sophisticated user to select a new SV. Bit 19 of the HOW is significant in that it indicates whether or not the leading edge of the TLM word is in synchronization with the X1 epoch that is required for a successful transfer from the C/A to P-code lock-on. Bits 20 through 22 contain subframe identification.

The remaining information in the navigation message is provided by the Control Segment and includes three blocks of data plus a block reserved for special messages. Block I data, which is contained in subframe 1, includes the clock correction parameters a_0, a_1, and a_2 and the presently unused parameters α_0, α_1, α_2, α_3, β_0, β_1,β_2, and β_3 used in the model for correction of atmospheric delays in propagation of the signal. Both the clock correction parameters and the ephemeris parameters that are contained in Data Block II are updated every hour. The Control Segment provides values for each hour in the 24 hours following update of the SV and the SV updates the navigation message each hour. Block I data also include 8 bits for the parameter T_{GD}, which permits an approximation of the atmospheric delay for receivers using only the L_1 signal and requiring less precision in their navigation. An 8-bit word, age of data (AODC), indicates the time since the last navigation upload for the use of more sophisticated receivers who may wish to select a satellite with a more recent update. The data block includes two spare words with 24 bits available in each.

(1)The P-code generated in each space vehicle is actually a seven-day-long phase segment of the P-code, which has a complete cycle of 267 days. The P-code is the modulo-2 sum of the output of two PRN code generators designated X_1 and X_2, each of which employs an input from the sum of the output of two subsidiary generators. All the space vehicles employ the same P-code generator, with each one assigned and generating a unique and mutually exclusive seven-day-long code-phase segment of the 267-day code. The P-code epoch (initial state) is made to occur in each space vehicle every seven days by resetting the X_1 and X_2 code generators to their initial states at the end of each week. The P signal generator in the j^{th} space vehicle is made to generate its unique phase segment of the P-code by offsetting (delaying) its X_2 generator j bits from its initial state at the time of reset. This technique places each X_2 register in the configuration existing at the beginning of its assigned seven-day phase segment of the P-code.

Data Block II, which contains the ephemeris prediction parameters, covers Subframes 2 and 3. This block also includes an age-of-data word (AODE) that indicates the time of the last measurement that was used to estimate the parameters and a spare 14 bits.

Subframe 4 is reserved for special messages which may be included in the data.

Block III data, which is located in Subframe 5, contains the "Almanac data." These data include information on the ephemerides, clock correction parameters, and atmospheric delay parameters for the normal complement of 24 satellites, plus one spare. The data are a subset of Block I and Block II parameters with reduced precision plus health and identification words for the space vehicles. The data are required to facilitate the rapid selection of four space vehicles for use in the navigation solution. The receiver uses the information first to identify which satellites are in view and to solve the algorithm that indicates the four that will provide the best navigation solution. The codes of these four are then generated by the receiver for matching with the corresponding codes among all the incoming signals. The almanac data also permit computation of the approximate range to each space vehicle and thereby facilitate the selection of the correct time slice in searching the codes for match.

The total Almanac data exceeds the capacity of the single Subframe 5 so that it is transmitted on a rotating page basis. The complete Almanac is contained in 25 frames. Sophisticated receivers will maintain almanacs in their data storage that preclude the need to wait for the transmission of the complete Almanac. The current information in their data storage is updated at the time of a GPS navigation fix. External almanacs, maintained from published data, may be used by operators with unsophisticated receivers.

The user algorithms for computing ephemeris position with almanac data are essentially the same as the user algorithms for computing the precise ephemeris from Block II parameters. The Block III data include the health word for each spacecraft to permit user rejection of space vehicles with unsatisfactory operation.

ERROR SOURCES

Errors contained in the pseudo-range measurements can be divided into the categories listed in Table 1. Various corrective techniques are employed in the system to reduce the magnitude of these errors in range that impact the estimates of user position. These are discussed below with the magnitude of the residual uncorrected errors summarized in Table 1. As discussed earlier, the range error resulting from the user clock bias is determined by the solution of the range equations so that it does not appear as an error in the ultimate range estimate.

Table 1. Range Error Budget

UNCORRECTED ERROR SOURCE	USER EQUIVALENT RANGE ERROR, 1σ	
	Meters	Feet
SV clock errors Ephemeris errors	1.5	5.0*
Atmospheric delays	2.4 - 5.2	8.0 - 17.0
Group delay (SV equipment)	1.0	3.3
Multipath	1.2 - 2.7	4.0 - 9.0
Receiver noise and resolution Vehicle dynamics	1.5	5.0
RSS	3.6 - 6.3	11.8 - 20.7
*Two hours after update		

Space Vehicle Clock Errors

Individual space vehicle clocks, although highly stable, may deviate as much as 976 microseconds from GPS system time. The offset is corrected by the model used by the receiver that employs the clock correction coefficients which are transmitted as data in the navigation message. The uncorrected errors due to clock deviation alone are very small (on the order of one foot of equivalent range); however, they are indistinguishable from certain components of ephemeris errors so that they are combined with the ephemeris errors in the error budget.

Atmospheric Delays

The time delay of RF signals passing through the ionosphere is due to a reduction in speed and the bending of the ray, both effects being due to refraction. The overall delay in the signal is nearly inversely proportional to the square of the frequency. The transmission of the navigation signal at the two frequencies (L_1 and L_2) is provided so that the magnitude of the delay can be calculated by a comparison of the two frequencies and removed with a satisfactory degree of accuracy. Forms of models are currently being evaluated during the Phase I program for users of both L_1 and L_2 and users of L_1 alone. These models employ the correction parameters ($\alpha_{0,1,2,3}$ and $\beta_{0,1,2,3}$) that are generated by the MCS and are uploaded periodically in the space vehicles for transmission to GPS users. Tropospheric errors are independent of frequency. They are relatively small but can be modeled fairly simply by receivers, using the elevation angle of the spacecraft. The combined effect of unmodeled ionospheric and tropospheric errors are estimated to result in a SV-to-user range error of from 2.44 to 5.18 meters (8 to 17 feet).

Group Delay

Group delay is defined as the delay resulting from uncertainties caused by the processing and passage of the signal through the SV equipment. The magnitude of these delays are calibrated during ground tests of the equipment. Corrections for the overall effect of these delays are included in the SV time offset correction parameters a_0, a_1, and a_2, discussed above. The estimated allowance for uncertainties in the group delay is 1 meter (3.28 feet).

Ephemeris Errors

Satellite ephemerides are determined by the MCS based on monitoring of individual space vehicle navigation signals by four monitoring stations. This operation results in a sort of inverted range process which enables the MCS to calculate the position of a space vehicle as if it were the user and the four monitoring stations were the space vehicles. The ephemeris determination is aided by precision clocks at the monitoring stations and by daily tracking over long periods of time with optimal filter processing. The determination process derives progressively refined information defining the gravitational field influencing the spacecraft motion; solar pressure parameters; and the locations, clock drifts, and signal delay characteristics of the monitoring stations. Based on this process, the MCS generates the ephemeris parameters that are uploaded in the SV periodically and included in the navigation data message where they are employed in the ephemeris model to calculate spacecraft position at the time of transmission of the received signals.

The satellite position errors resulting from this process are still on the order of several meters. However, it is the ranging errors that are of primary significance, and the effect of these is relatively small. Errors that are common to the four range measurements, for example, will cause an apparent error in the user's clock, which tends to be compensated for in calculation of the user clock bias. The combined effects of residual uncertainties in SV clock offsets and ephemeris determinations are estimated to result in range errors of about 1.5 meters (5 feet).

Multipath

Multipath errors result from the combination of data from more than one propagation path that distorts the signal characteristics from which the range measurements are made. These errors are dependent on the nature and location of reflective surfaces peculiar to each user location. Aircraft simulations have indicated these errors to be relatively small and have resulted in an error estimate of from 1.2 to 2.7 meters (4.0 to 9.0 feet).

Receiver Noise and Resolution

Noise and resolution errors resulting from the processing of signals by the receiver hardware and software will contribute to errors in the determination of range. With high-performance four-channel receivers, it is expected that these errors will be about 1.52 meters (5 feet).

Receiver Vehicle Dynamics

User vehicle dynamics will contribute to the total ranging errors. These can be compensated for by special receiver designs and by Kalman optimal processing of received signals, which would be required, for example, by high-speed, very low altitude aircraft. The estimated error due to receiver noise and resolution, above, is based on nominal receiver vehicle dynamics. No allowance is made in the overall error budget for high vehicle dynamics.

The overall SV-to-user range uncertainties due to the combined error sources identified in Table 1 are estimated to be from 3.6 to 6.3 meters (12 to 21 feet). The uncertainties in range to the space vehicles combine with the geometry of the SV positions in their effect on the accuracy of user position estimates.

SYSTEM ACCURACY

Geometric Dilution of Precision (GDOP)

The magnitude of the ranging errors, combined with the geometry of the four selected satellites, will determine the magnitude of the user position errors in the GPS navigation fix. The effect of geometry is expressed by the geometric dilution of precision (GDOP) parameters. The use of GDOP was originally developed in connection with LORAN navigation systems. Extended to the GPS system, with fixes in three dimensions plus time, the parameters include PDOP, which reflects the dilution of precision in position in three dimensions; HDOP, dilution of precision in the two horizontal dimensions; VDOP, dilution of precision in the vertical dimension; and TDOP, dilution of precision in time, i.e., in the estimate of the range equivalent of the user clock bias.

PDOP	x	User to Satellite Range Error, 1σ	=	Radial error in user position 1σ, in 3 dimensions
HDOP	x	User to Satellite Range Error, 1σ	=	Radial error in user position, 1σ, in the horizontal plane
VDOP	x	User to Satellite Range Error, 1σ	=	Vertical error in user position, 1σ
TDOP	x	User to Satellite Range Error, 1σ	=	Error, 1σ, in the range equivalent of the user clock offset

Hence, small values of the GDOP parameters indicate good arrangements in the geometry of the selected satellites and correspondingly small errors in position and time fixes. Figure 4 gives the values of PDOP for cumulative proportions of users evenly distributed over the globe and around the clock who select the best four satellites from those that are visible 5 degrees or more above the horizon. The rms value of PDOP is in the neighborhood of 2.60 which, when combined with the range errors taken from Table 1 of from 3.6 to 6.3 meters (11.8 to 20.7 feet), gives user three-dimensional position errors of from 9.4 to 16.4 meters (31 to 54 feet) 1σ. The horizontal component of the position error reflected by HDOP is usually less. The corresponding rms value of HDOP is about 1.45, which yields horizontal position errors of from 5.2 to 9.1 meters (17 to 30 feet) 1σ. The corresponding rms value of TDOP is about 1.2 which, when range is converted to time, yields a 1σ time error of from 14 to 25 nanoseconds.

The value of GDOP itself is a composite measure that reflects the influence of satellite geometry on the combined accuracy of the estimate of user time (user clock offset) and user position.

$$GDOP = \sqrt{(PDOP)^2 + (TDOP)^2}$$

The four "best" satellites selected by the user receivers are those with the lowest GDOP. A high correlation between GDOP and the volume of the tetrahedron formed by the points of unit vectors from the user to the satellites has provided a relatively simple algorithm that can be used by receivers for the selection. A derivation of GDOP is given below in the section with that heading.

User Error Distribution

The GDOP parameters are generally given in terms of percentiles, such as the fiftieth and ninetieth percentiles, because the values of the GDOP parameters are statistically distributed in a non-Gaussian fashion that tends to distort average or mean square values. The lack of a mathematical expression for the distribution of the GDOP parameters, however, has prevented the establishment of a mathematical relationship between percentiles of GDOP and percentiles of user navigation errors; i.e., there is not a correspondence between the fiftieth or ninetieth percentile of GDOP parameter and the fiftieth or ninetieth percentile of user navigation error.

Paul Jorgensen of Aerospace Corporation has addressed the problem of determining the distribution of user navigation errors by conducting a Monte Carlo simulation that has provided a good approximation.(1) The simulation was based on a trial for several thousand user observations evenly distributed over the globe and in time, using the NAVSTAR twenty-four satellite baseline constellation. Calculated GDOP parameters for the best four satellites in each case were combined with random selections from a Gaussian distribution representing pseudo-range errors to calculate the user navigation and time errors. The range errors were normalized with zero expected means and 1σ values equal to unity. Their distribution can justifiably be assumed to be Gaussian because they are a composite of a large number of generally independent error sources. The results have been provided in the form of normalized user error distribution parameters. When mean, rms, or percentile values of these parameters are multiplied by 1σ range errors, they yield corresponding values of the user navigation error and range equivalent time error distributions. Values of the parameters are given in Table 2, along with corresponding values of the user errors based on the 1σ range error estimation of from 3.6 to 6.3 meters (11.8 to 20.7 feet) given in Table 1. Also shown are the corresponding values of GDOP parameters which indicate the smaller spread of their distributions.

Velocity Measurement with NAVSTAR

Although the NAVSTAR system is designed primarily for precision position estimates, it will also provide precision velocity measurements, which will be of interest to military and other classes of users. The velocity estimates are made by measuring the doppler shift in the carrier frequency of the navigation signal from the satellites. Precision measurements are possible because of the precise knowledge of the satellite ephemerides and because of the short wavelength of the carrier frequency, which is approximately 19 cm. In addition, the error offset in the frequency of the receiver oscillator can be solved by the use of four satellites and four range rate equations in a manner analogous to the solution for the user clock offset.

The effect of satellite geometry on the relationship between range rate error and user velocity errors is completely analogous to that between range errors and user position errors so that the GDOP parameters, satellite selection algorithm, and the user navigation error parameters described above are equally applicable to the velocity measurements.

---

(1) P.S. Jorgensen, "Normalized Accuracy of the NAVSTAR/Global Positioning System," The Aerospace Corporation, Report No. TOR-0078(3475-10)-2 (28 February 1978).

Table 2. Anticipated Worldwide User Position Error Distribution

	HORIZONTAL				VERTICAL				TIME		
	HDOP	User Error Parameter	User Error (Meters)	User Error (Feet)	VDOP	User Error Parameter	User Error (Meters)	User Error (Feet)	TDOP	User Error Parameter	User Error (Nano-seconds)
50th percentile	1.39	1.15	4.1-7.2	14-24	1.99	1.39	5.0-8.8	16-29	1.05	0.73	8-15
rms	1.44	1.45	5.2-9.1	17-30	2.16	2.21	8.0-13.9	26-46	1.21	1.22	14-25
90th percentile	1.71	2.19	7.9-13.8	26-45	2.80	3.57	12.9-22.5	42-74	1.76	1.96	23-40

Based on Range Error Budget 11.8-20.7 feet; 24-Satellite Baseline Constellation; 5-degree Satellite Elevation Mask Angle

Receiver dynamics will have a major impact on the accuracy of velocity measurement which can be attained. It is anticipated that other error sources will be small. Receiver dynamics can impact the measurements both by introducing noise in the phase-lock tracking loop and by affecting the oscillator frequency.

High quality receivers in a benign environment with averaging intervals on the order of a second will be able to make user-to-satellite range rate measurements with accuracies of a few hundredths of a foot per second, 1σ. Accuracies of 0.061 to 0.15 mps (0.2 to 0.5 fps) are anticipated, with high quality receivers, in a severe dynamic environment with loadings up to the order of 5 g's. Table 3 gives the estimated magnitude of user velocity errors based on range rate errors of 0.015 and 0.061 mps (0.05 and 0.2 fps) using the navigation measurement error parameters described in the foregoing section.

Table 3. World-Wide User Velocity Measurement Errors

Distribution Measure	Based on 0.015 mps (0.05 fps) Range Rate Errors				Based on 0.061 mps (0.2 fps) Range Rate Errors			
	Horizontal Velocity Error		Vertical Velocity Error		Horizontal Velocity Error		Vertical Velocity Error	
	(mps)	(fps)	(mps)	(fps)	(mps)	(fps)	(mps)	(fps)
50th percentile	.02	.06	.02	.07	.07	.2	.08	.3
rms	.02	.07	.03	.11	.09	.3	.13	.4
90th percentile	.03	.11	.05	.18	.13	.4	.21	.7

THE GPS NAVIGATION SOLUTION

The GPS navigation solution can be implemented by the use of vectors and matrix algebra, with the initial range equations set up as follows. (See Figure 5.)

$$\bar{R}_u = \bar{R}_i - \bar{D}_i \tag{1}$$

where

$\bar{R}_u$ = the vector from the center of the earth to the user

$\bar{D}_i$ = the vector from the user to the ith satellite

$\bar{R}_i$ = the vector from the center of the earth to the ith satellite

i = 1 to n satellites ≥ 4

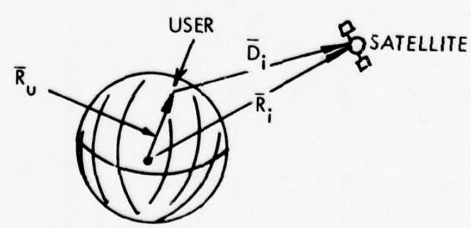


Figure 5. Navigation Solution Geometry

A solution for the vector $\bar{R}_u$ will yield the desired three unknowns of user position, identified as R_{ui} in a three-axes reference coördinate system. The components of $\bar{R}_i$ are known based on the satellite ephemerides. $\bar{D}_i$ must be dealt with in terms of its magnitude, which is determined by the measured value of pseudo-range with allowances for user and satellite clock corrections.

By defining $\bar{e}_i$ as the unit vector from user to satellite and noting that $\bar{e}_i \bullet \bar{D}_i = D_i$ (the magnitude of $\bar{D}_i$), (1) becomes

$$\bar{e}_i \bullet \bar{R}_u = \bar{e}_i \bullet \bar{R}_i - D_i \tag{2}$$

The range, D_i, to the satellite can be expressed as

$$D_i = \rho_i - B_u - B_i, \tag{3}$$

where ρ_i is the measured pseudo-range. B_u and B_i, respectively, are the range equivalents of the user and satellite clock offsets.

Then (3) combined with (2) gives

$$\bar{e}_i \bullet \bar{R}_u - B_u = \bar{e}_i \bullet \bar{R}_i - \rho_i + B_i \tag{4}$$

The set of equations (4) are the basic range equations which contain four unknowns, consisting of the three axis components of user position in $\bar{R}_u$ and the range equivalent of the user clock offset, B_u, and therefore require four equations for solution. A solution can be provided employing the following matrix definitions:

$X_{u(4x1)} \triangleq \left| R_{u1}, R_{u2}, R_{u3}, -B_u \right|^T$, which contain the unknowns of user position and clock correction.

$$G_{u\,(n\times 4)} \triangleq \begin{vmatrix} \Gamma_1 \\ \Gamma_2 \\ \Gamma_3 \\ \vdots \\ \Gamma_n \end{vmatrix} \qquad A_{u(n\times 4n)} \triangleq \begin{vmatrix} \Gamma_1 & O & O \cdots & O \\ O & \Gamma_2 & O & O \\ O & O & \Gamma_3 & O \\ \vdots & \vdots & \ddots & \vdots \\ O & O & O \cdots & \Gamma_n \end{vmatrix}$$

WHERE, $\Gamma_i \triangleq (e_{i1}, e_{i2}, e_{i3}, 1)$

$O \triangleq (0, 0, 0, 0)$

Note that e_{ij} are the components of the unit vectors, $\bar{e}_i$, in each of the three axes and are the direction cosines from the user to the satellites.

$$\bar{S}\ (4n\text{x}1) \triangleq \left| R_{11}, R_{12}, R_{13}, B_1, R_{21}, R_{22}, R_{23}, B_2 \ldots \right.$$

$$\left. \ldots R_{n1}, R_{n2}, R_{n3}, B_n \right|^T$$

$$\bar{\rho}_{(n\text{x}1)} \triangleq \left| \rho_1, \rho_2, \rho_3, \ldots \rho_n \right|^T$$

The following arrangement of the defined matrices represents the set of equations (4) and will facilitate the desired solution for $\bar{X}_u$:

$$G_u \bar{X}_u = A_u \bar{S} - \bar{\rho} \tag{5}$$

A general least squares solution usable with any number of satellites ≥ 4 may be had by factoring G^T into both sides of (5) giving

$$G_u^T G_u \bar{X}_u = G_u^T [A_u \bar{S} - \bar{\rho}].$$

Then,

$$\hat{X}_u = [G_u^T G_u]^{-1} G_u^T [A_u \bar{S} - \bar{\rho}], \tag{6}$$

which provides the desired solution for $\bar{X}_u$. Note that the G_u and A_u matrices are made up primarily of e_{ij}, which are the direction cosines from the user to the satellites. The solution of (6) requires an iterative solution based on initial estimates of the direction cosines e_{ij} made from an independent estimate of user position.

Derivation of GDOP

The covariance matrix of the error, $\delta \bar{X}_u$, in the estimate of $\bar{X}_u$ is given by

$$\mathrm{Cov}\,\delta \bar{X}_u = (G_u^T G_u)^{-1} G_u^T \,\mathrm{Cov}\,\delta\, [A_u \bar{S} - \bar{\rho}][(G_u^T G_u)^{-1} G_u^T]^T \tag{7}$$

This gives the covariance of the errors in user position and user time ($\delta \bar{X}_u$) if the measurement error statistics represented by Cov $\delta(A_u \bar{S} - \bar{\rho})$ are accurately known. GDOP is calculated by setting Cov $\delta(A_u \bar{S} - \bar{\rho})$ equal to the identify matrix. The remaining portions of (7) can then be reduced to the following:

$$\mathrm{Cov}\,\delta \bar{X}_u = (G_u^T G_u)^{-1} \tag{8}$$

Cov $\delta(A_u \bar{S} - \bar{\rho})$ primarily reflects range measurement error statistics (e.g., satellite ephemeris, ionospheric model, and instrumentation errors), whereas, G_u reflects only the geometry of the system. A good approximation to the effect of geometry is therefore measured when it is assumed that (Cov δ $(A_u \bar{S} - \bar{\rho})$ equals the identity matrix. This assumption normalizes the relationship by setting the range errors equal to one, with zero mean, and implies that the range errors are both equal and independent.

Cov $\delta \bar{X}_u$ appears as,

$$(G_u^T G_u)^{-1} = \begin{array}{c c} & \begin{array}{cccc} X & Y & Z & \mathrm{Time} \end{array} \\ \begin{array}{c} X \\ Y \\ Z \\ \mathrm{Time} \end{array} & \begin{bmatrix} \sigma^2_{xx} & \sigma^2_{xy} & \sigma^2_{xz} & \sigma^2_{xt} \\ \sigma^2_{yx} & \sigma^2_{yy} & \sigma^2_{yz} & \sigma^2_{yt} \\ \sigma^2_{zx} & \sigma^2_{zy} & \sigma^2_{zz} & \sigma^2_{zt} \\ \sigma^2_{tx} & \sigma^2_{ty} & \sigma^2_{tz} & \sigma^2_{tt} \end{bmatrix} \end{array}$$

where the diagonal values are the variance of the estimated user position in each axis and in the user time offset. The GDOP factors are defined as

$$\mathrm{HDOP} = \sqrt{\sigma^2_{xx} + \sigma^2_{yy}}$$

$$\mathrm{VDOP} = \sigma_{zz}$$

$$\mathrm{PDOP} = \sqrt{\sigma^2_{xx} + \sigma^2_{yy} + \sigma^2_{zz}}$$

$$\mathrm{TDOP} = \sigma_{tt}$$

$$\mathrm{GDOP} = \sqrt{\sigma^2_{xx} + \sigma^2_{yy} + \sigma^2_{zz} + \sigma^2_{tt}}$$

The product of the GDOP factors and estimates of the errors in the range measurements then give an estimate of the corresponding errors in user position or in user time.

AD-A075 208 ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT--ETC F/G 17/7
PRINCIPLES AND OPERATIONAL ASPECTS OF PRECISION POSITION DETERM--ETC(U)
JUL 79 C T LEONDE
UNCLASSIFIED AGARD-AG-245 NL

AD
A075208

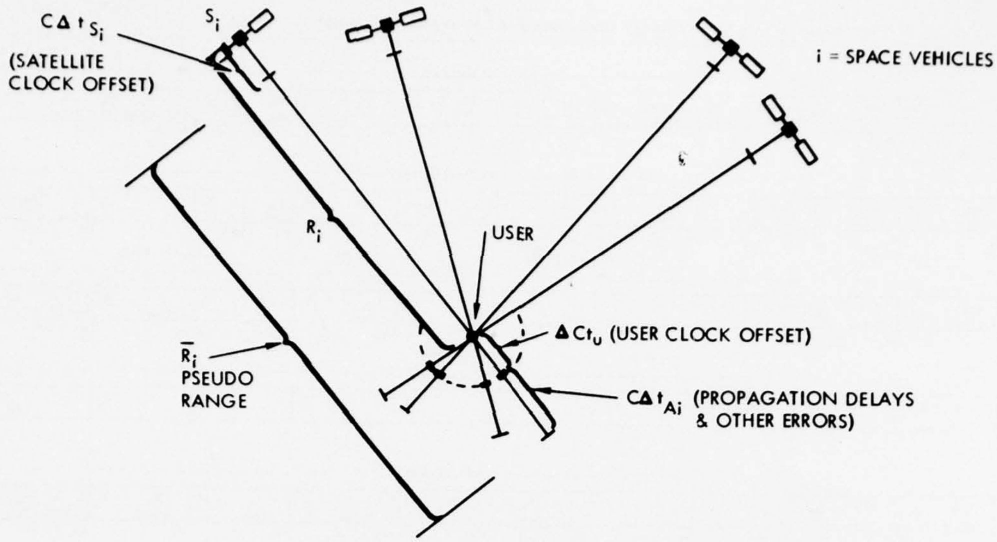


Figure 1. Pseudo - Range

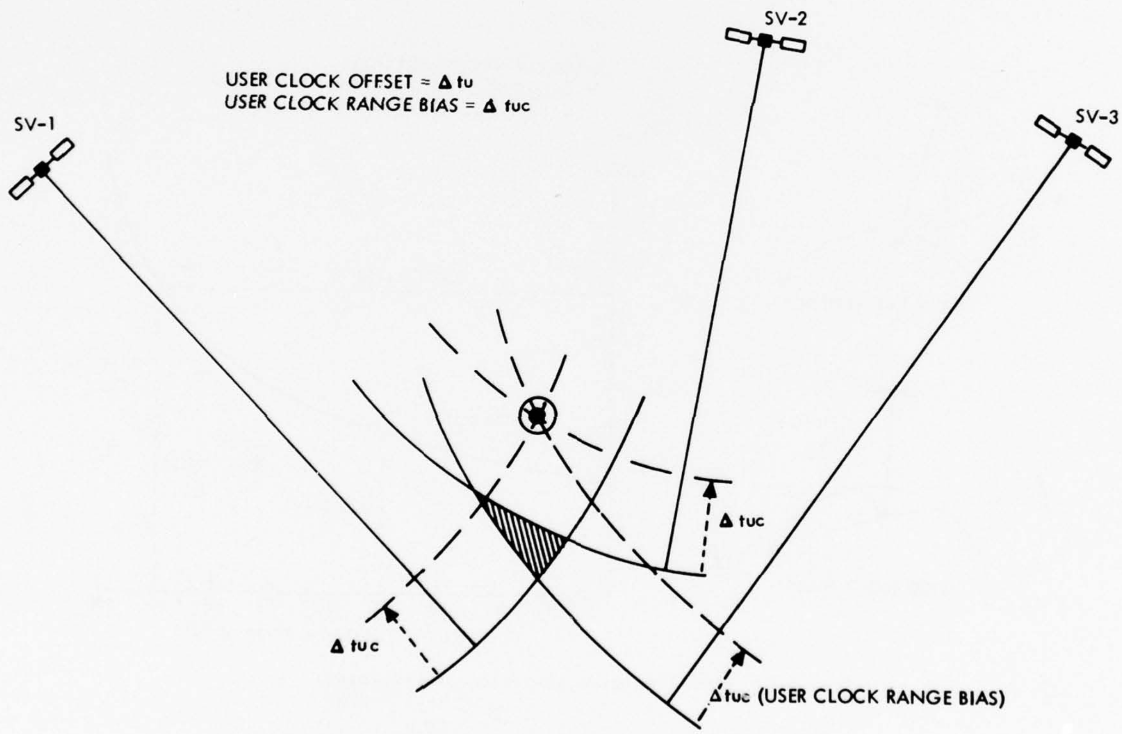
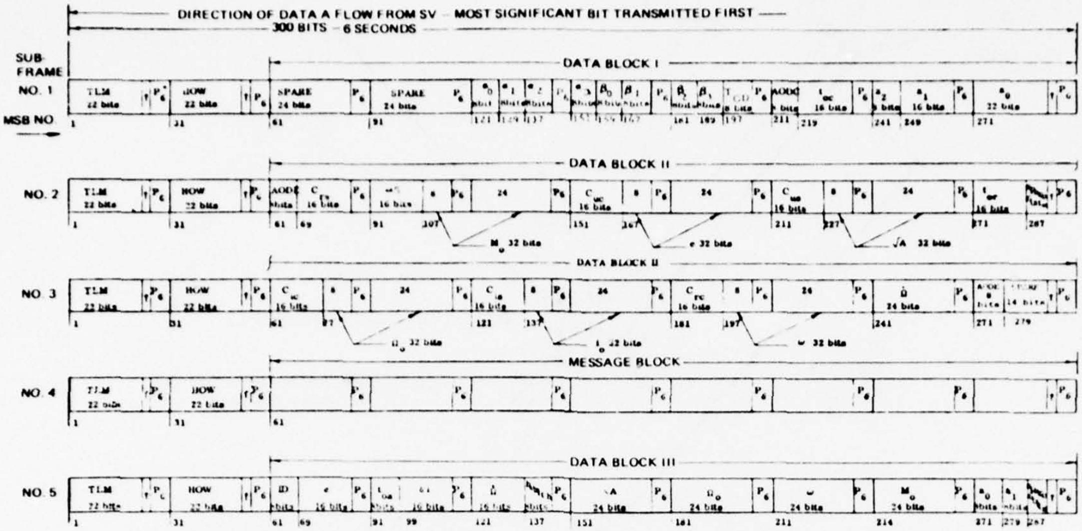


Figure 2. Determination of User Position and Time Offset
(Two Dimensional Case)



* P_6 – 6 parity bits

† – 2 noninformation bearing bits used for parity computation (see para. 3.3.2.10.6)

NOTE: All binary numbers are two's complement.

Figure 3. Signal Data Frame Content

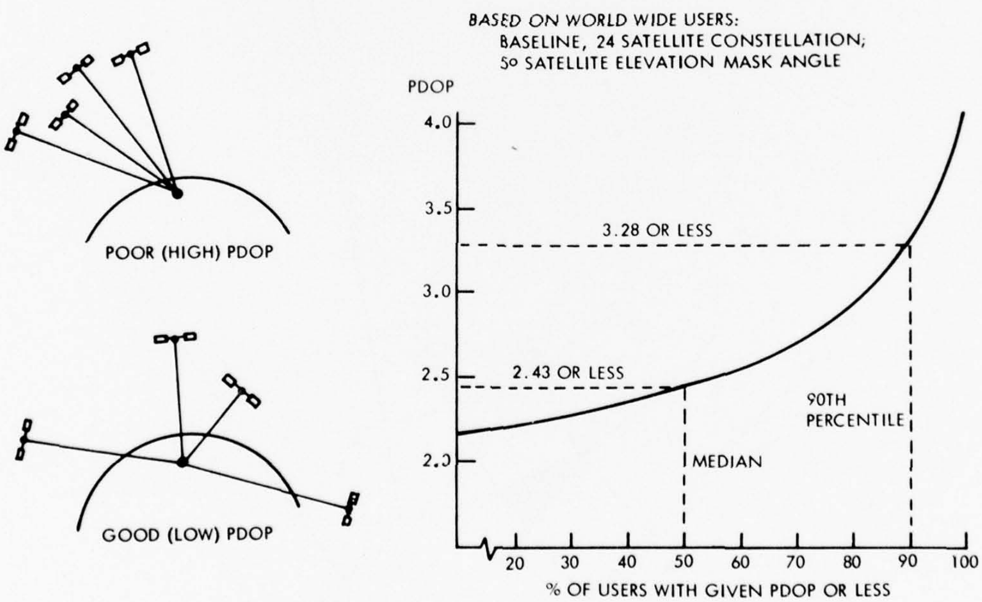


Figure 4. PDOP (Position, Dilution of Precision)

GLOBAL POSITIONING SYSTEM: SIGNAL STRUCTURE AND PERFORMANCE CHARACTERISTICS

James J. Spilker, Jr.
Stanford Telecommunications, Inc.
1195 Bordeaux Drive
Sunnyvale, California 94086

SUMMARY

Details of the GPS signal structure are discussed as relates to the signal generation and the performance of the navigation system. GPS performance objectives, orbit geometry, and propagation effects are summarized in order to gain better understanding of the signal and what characteristics it must provide. With these performance objectives as a preface, the details of the signal are described, showing the details of the dual frequency transmission and both the precise (P) and clear/acquisition (C/A) codes and their characteristics. Finally, the basic performance of simplified receivers operating on this received signal in order to show compatibility with the original performance objectives and typical receiver operation. It is shown that an rms position error of less than 10 meters is well within the achievable performance bounds of the system.

SECTION 1

INTRODUCTION AND PERFORMANCE OBJECTIVES

1.1 Introduction

In this paper we describe the detailed signal structure used in the Global Positioning System satellite navigation system. In order for one to understand the performance characteristics in a meaningful sense, we begin by discussing, in an idealized sense, the concepts for high accuracy, real-time navigation using satellites. The various perturbation effects on the navigation signal and overall system are then described. These perturbations include relativistic effects, multiple access interference between satellites, tropospheric and ionospheric propagation delays, multipath, thermal noise, and other interference effects. We conclude the first section with a summary of the performance objectives for the signal.

The paper then continues with a detailed discussion of the signal structure, the code properties, and the performance of the signal relative to the various objectives and constraints in the first section.

In the concluding section the performance capability of a typical receiver for this signal is described and briefly analyzed. The search, acquisition, and tracking accuracy for the GPS codes are included. Effects of user dynamics are considered. Multipath and other interference error effects are summarized.

1.2 Performance Objectives

There are several key performance objectives for the GPS system which distinguish it from previous satellite and land-based navigation systems. Some of the more important are summarized below:

- High Accuracy 10-30 meter rms position error
- Real-Time navigation for users with high dynamics
- World-Wide Operation
- Tolerant to Nonintentional or Intentional Interference

In addition to these constraints, the user cannot be required to carry a precision atomic clock, and an initial navigation fix should be obtained within a reasonable period (minutes rather than hours) after initial turn-on of the receiver.

Obtaining high accuracy in real-time without ambiguity requires a relatively large bandwidth ($\geq$ 10 MHz) and a signal with a long period. The high dynamics of the user particularly in an aircraft requires the use of an omni-directional or hemispherical pattern antenna. Coupling these two requirements together leads to an RF frequency selection which is large compared to the bandwidth but not so large as to give too great a space loss. Space loss is computed with a combination of an earth coverage satellite antenna and a 0 dBI gain receive antenna. The RF frequency must also be consistent with available frequency allocations. L-band is selected.

The demand for world-wide operation primarily places a constraint on the satellite orbits. One must employ at least some satellites in inclined orbits in order to provide coverage to the polar regions. Secondly, the signals received from the satellite are of relatively low power. Hence the signal should be tolerant to low level interference which might, for example, be nothing more than a spurious harmonic of some lower frequency narrow-band signal.

1.3 Satellite Navigation Concepts

As an elementary example of the use of satellites for navigation, examine the single satellite in Fig. 1-1. The satellite carries a display of the satellite on-board clock

and position $\vec{x}_1 \triangleq (x_1, y_1, z_1)$ in earth-center-earth-fixed coordinates. If this hypothetical display could be viewed by an observer on earth through a telescope, then simultaneous photographs can be taken of the satellite clock, through the telescope, and a local clock. Both satellite and local clocks are assumed to be precisely at GPS system time. The time difference between the two clocks as observed by the user is exactly the propagation delay (neglecting relativistic effects).

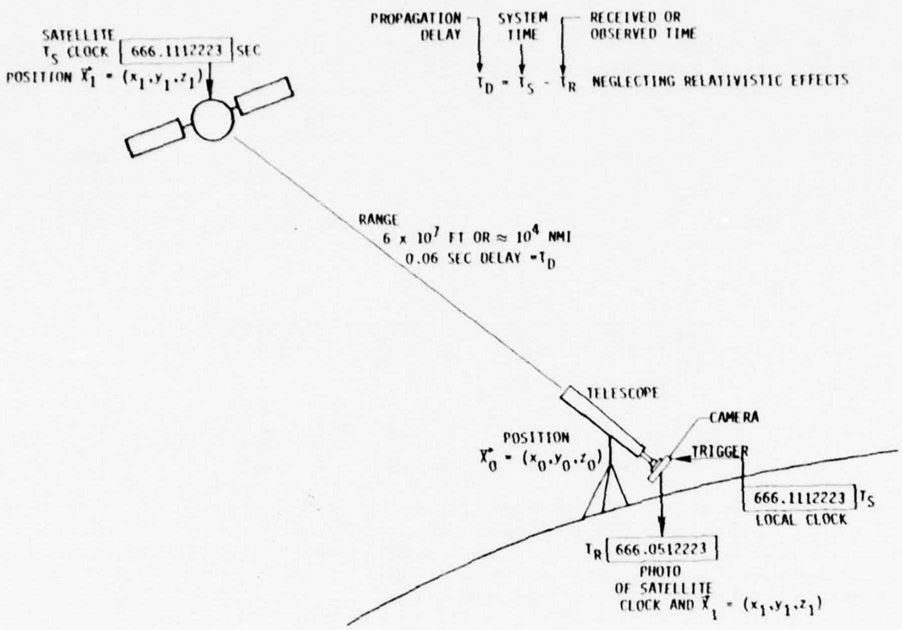


FIGURE 1-1 SATELLITE & USER CLOCK TIMING CONCEPTS - SIMULTANEOUS PHOTOGRAPHS OF CLOCKS

From this measurement and the knowledge of the satellite position one can of course determine that the user position is on the surface of a sphere centered at the satellite. Clearly if one takes a sequence of measurements and forms a sequence of spheres one can determine position from only one satellite. However, two of our performance objectives then have been violated: the solution is not in real-time, and an accurate clock has been assumed at the user.

The system configuration of Fig. 1-2 eliminates the requirement for a local precision clock and provides a real-time position measurement through the use of 4 measurements, and 4 equations to solve for the 4 unknowns, x_0, y_0, z_0, and T_S. One simultaneously photographs the 4 satellite displays to obtain T_A, T_B, T_C, T_D, and $\vec{X}_A$, $\vec{X}_B$, $\vec{X}_C$, $\vec{X}_D$. In addition one can measure the derivatives of these quantities and solve for user velocity.

In order to be useful, these simplified models of a navigation system must first be transformed into a realizable form. The clock display must be transformed into an equivalent RF signal which carries with it the time-of-day with sufficient precision and lack of ambiguity; the position of the satellite and any errors in the satellite clock must be carried to the user in a down-link data stream. Even an atomic standard time clock has some inaccuracy relative to system time (a set of atomic clocks). See Fig. 1-3 for a typical plot of satellite clock count vs "true" system time showing the clock time error.*

It is also easily shown that if the user has a good crystal clock that solutions of the above equations for user time and doppler can correct the user clock, and this correction will be valid over a reasonable period thereafter, depending on the crystal clock stability. For that time period one needs only to solve for three unknowns, and hence only needs 3 satellites in view.

1.4 GPS Orbit Configuration and Multiple Access

Figure 1-4 shows the GPS satellite configuration of 24 satellites. There are 3 orbit planes, each inclined by 63° with respect to the equatorial plane and offset from one another by 120° in longitude. Eight satellites are in circular prograde 12 hour** orbits in each orbit plane. The satellite altitude is approximately 16,020 Km or 10,898 nm. Figure 1-5 shows the orbit traces of the satellites in each orbit plane and the relative positions of each satellite in the 6 satellite Phase I configuration.

*J. J. Spilker, Jr., Digital Communications by Satellite, Prentice-Hall, Engelwood Cliffs, N.J., Chap. 17, description of precise measures of time.

**The 12 hr is in sidereal time, not solar time. Thus the subsatellite point at noon slowly shifts each day since a sidereal day is roughly 4 minutes shorter than a solar day.

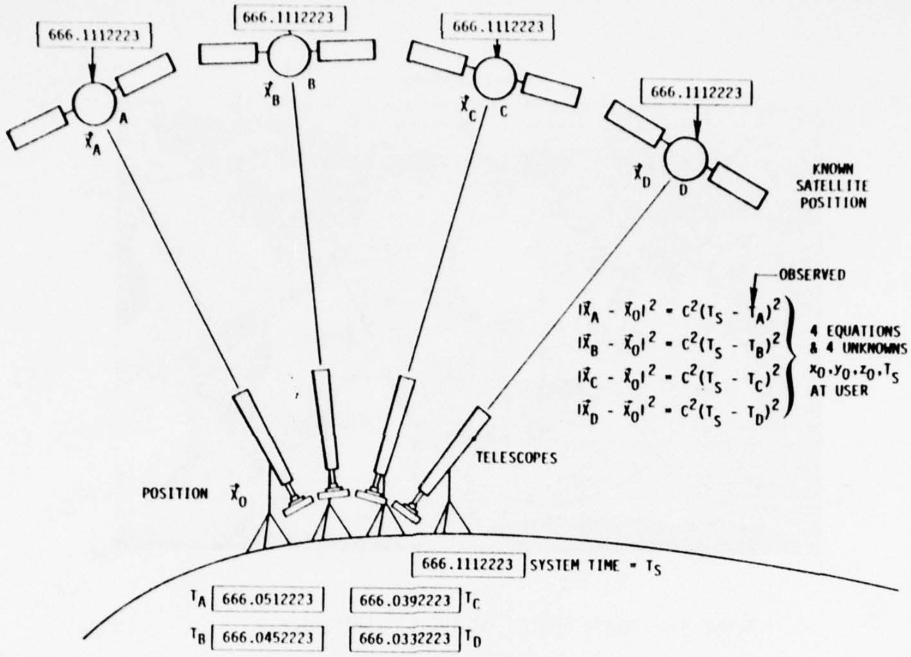


FIGURE 1-2 MULTI-SATELLITE CLOCK TIMING AND POSITION MEASUREMENT

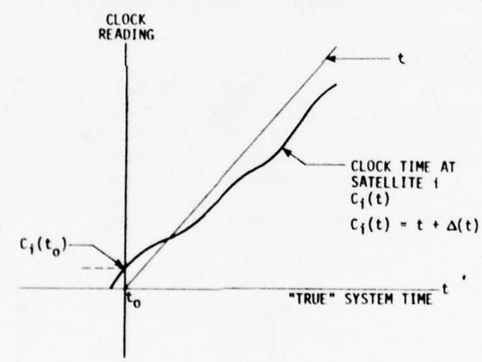


FIG 1-3 SATELLITE CLOCK TIME CHARACTERISTICS

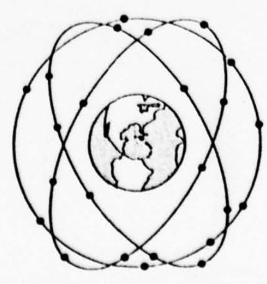


FIGURE 1-4 GPS ORBIT CONFIGURATION. THE SATELLITE ALTITUDE IS APPROXIMATELY 16,020 km OR 10,898 nm.

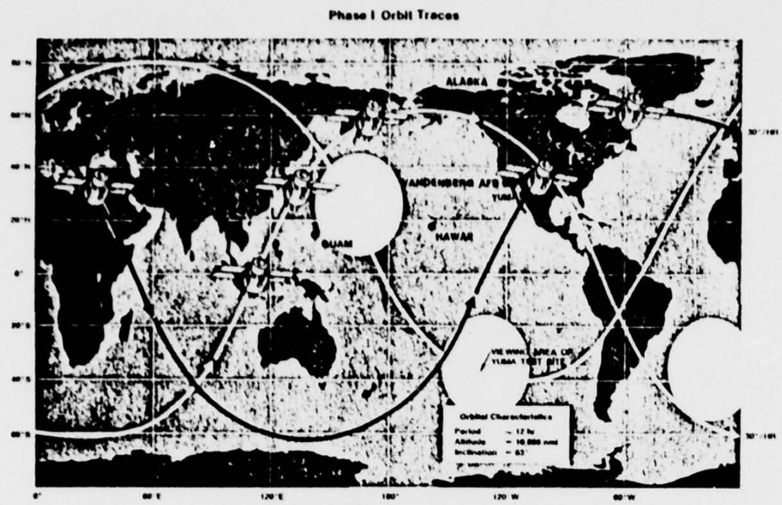


FIGURE 1-5 ORBIT TRACES OF THE GPS SATELLITE

In the fully implemented 24 satellite system there are always at least 6 satellites in view and as many as 11 (depending on user location). One required characteristic of the GPS signals then is that one must be able to observe these multiple satellite signals simultaneously without mutual interference. This property is analogous to multiple access in satellite communications.

Path Delays and Doppler Shifts

Closer examination of the signals received by a given user from multiple satellites illustrates the fact that these signals have different path delays and different doppler shifts (see Fig. 1-6). In Fig. 1-7 the maximum doppler shift (ft/sec) is shown as a function of user position for stationary users. The maximum doppler shift observed by a stationary observer is on the order of 2700 ft/sec or a $v/c = 2.7 \times 10^{-6}$ where v is the radial velocity to the user and c is the speed of light.

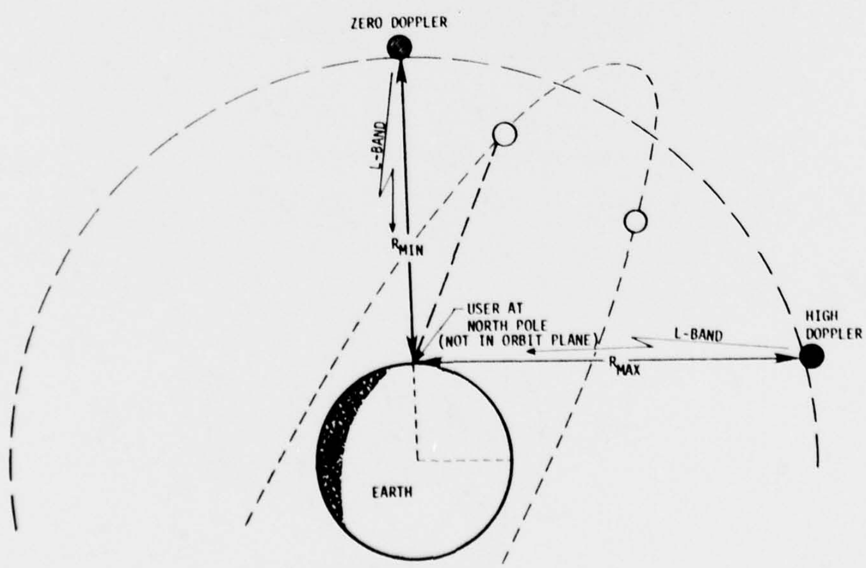


FIGURE 1-6 GPS RANGE AND DOPPLER FOR A USER AT THE POLE
RANGE VARIATION ≤ 3444 nm , MAX RANGE VARIATION IF USER IS IN ORBIT PLANE
DOPPLER VARIATION ≤ ±2654 ft/sec
1 nm = 1.852 km

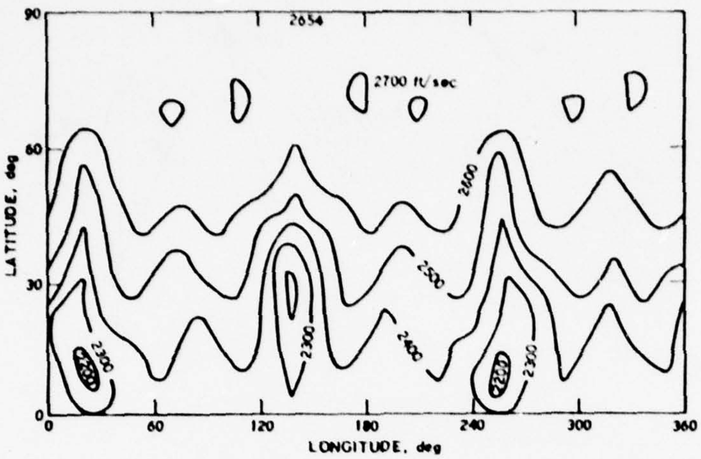


FIGURE 1-7 GPS PHASE III - 3x8 BASELINE, MAXIMUM RANGE RATE (DOPPLER)

If a stationary user is at the North Pole (zero earth rotation effect), the receiver can observe a +2600 ft/sec velocity from one satellite and a -2600 ft/sec velocity from another satellite nearly simultaneously. Thus the differential radial velocity can exceed 5000 ft/sec. For a 1.56 GHz-signal, this differential corresponds to 7500 Hz.

Geometric Dilution of Precision (GDOP)

The accuracy with which one can measure position and time is related to the accuracy in radial range measurement by factors known as the GDOP or Geometric Dilution of Precision. The rms position error

$$\sigma_p \triangleq \sqrt{\sigma_x^2 + \sigma_y^2 + \sigma_z^2} \tag{1-1}$$

is related to the rms radial range error σ_r by

$$\frac{\sigma_p}{\sigma_r} \triangleq \text{PDOP} \tag{1-2}$$

PDOP is the Position Dilution of Precision, and we similarly define $\sigma_h \triangleq \sqrt{\sigma_x^2 + \sigma_y^2}$ and the Horizontal Dilution of Precision is

$$\frac{\sigma_h}{\sigma_r} \triangleq \text{HDOP} \tag{1-3}$$

The value of PDOP can be determined geometrically by relating it to the volume of a special tetrahedron as shown in Fig. 1-8. The user position is at point $\bar{P}$, and the satellites are located at points $\bar{R}_n$. Generate a unit sphere centered at the user $\bar{P}$ and draw vectors intersecting the sphere to each satellite as shown in Fig. 1-8(b) where one of the satellites is shown at the user's zenith. The tetrahedron formed by connecting these points together and the user point $\bar{P}$ has a volume V. It can be shown that

$$\text{PDOP} \sim 1/V \tag{1-4}$$

Thus as the volume of the tetrahedron becomes larger, the PDOP becomes smaller, and hence the position accuracy improves. The volume is maximized when the one satellite is at the user's zenith and the other three are separated by 120° and are as low on the horizon as permitted by the user's antenna elevation angle (maximize the horizontal cross-sectional area).

Figure 1-9 shows the various GDOP factors vs the cumulative probability of achieving a given GDOP or lower. The values shown are for a 5° elevation mask, i.e. only satellites in the 24 satellite constellation above 5° elevation angle are assumed to be in view. Clearly one has a high probability of a PDOP of 3 or less. Thus if one is to have a 10 meter accuracy goal the desired accuracy in range measurement should be on the order of (1/3) 10 meters or roughly $\sigma_r < 10$ nsec.

If one assumes that a modulation component or code chip modulating the signal can be resolved to 1-10% of its width with reasonable ease then the 10 nsec value for σ_r leads to

a code chip width of 1 μsec to 100 nsec or a 1-10 Mbps code clock rate.

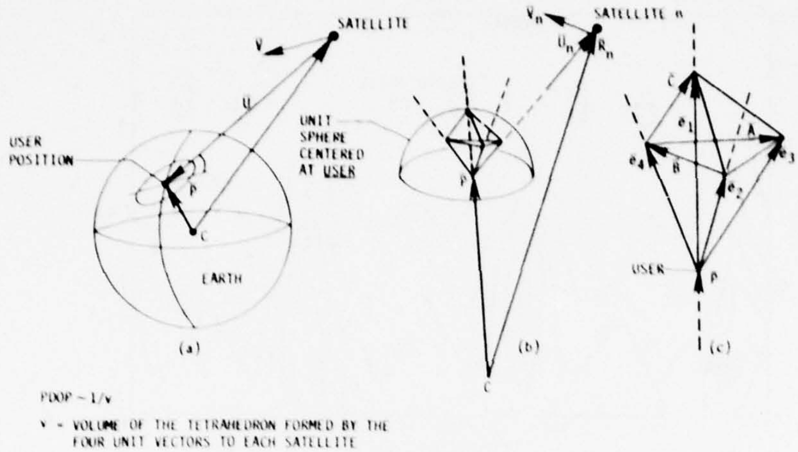


FIGURE 1-8

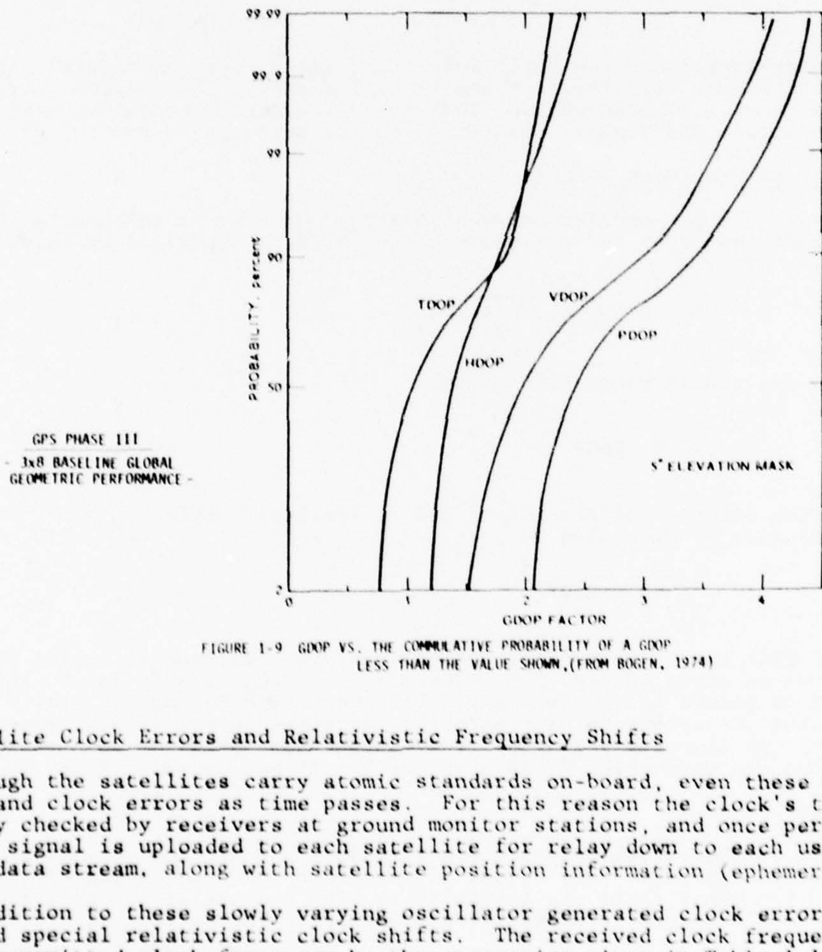


FIGURE 1-9 GDOP VS. THE COMMULATIVE PROBABILITY OF A GDOP LESS THAN THE VALUE SHOWN.(FROM BOGEN, 1974)

1.5 Satellite Clock Errors and Relativistic Frequency Shifts

Although the satellites carry atomic standards on-board, even these clocks are subject to drifts and clock errors as time passes. For this reason the clock's timing errors are continually checked by receivers at ground monitor stations, and once per day, a clock correction signal is uploaded to each satellite for relay down to each user as part of the satellite data stream, along with satellite position information (ephemeris).

In addition to these slowly varying oscillator generated clock errors there are also general and special relativistic clock shifts. The received clock frequency f_r differs from the transmitted clock frequency by the expression given in Table 1-1. These are two types of effects. The first is caused by the difference in gravitational potential between the satellite and the user. The second effect is caused by the difference in velocities of the users. Both of these effects depend somewhat on where the user is on the earth. The average effect is a net fractional increase in frequency of 447.9×10^{-12}. The effect of the sun causes a minor perturbation in this value, and the moon a still smaller effect.

Much of this relativistic effect can be corrected by purposely setting the satellite clock frequency slightly low by the factor 4.45×10^{-10}. The satellite signal speeds up as it approaches earth causing the observed frequency to increase. The remaining correction

is carried in the downlink data stream.

TABLE 1-1 RELATIVISTIC CLOCK FREQUENCY SHIFTS

- RECEIVING STATION CLOCK FREQUENCY f_r
- TRANSMITTING SATELLITE CLOCK FREQUENCY f_t

$$\frac{f_r}{f_t} = 1 + \underbrace{\frac{1}{c^2}\left(\phi_t - \phi_r\right)}_{\substack{\text{GRAVITATIONAL POTENTIAL} \\ \text{(General Relativity)}}} + \underbrace{\frac{1}{2}\left(\frac{V_r^2}{c^2} - \frac{V_t^2}{c^2}\right)}_{\substack{\text{SPECIAL RELA-} \\ \text{TIVISTIC SHIFTS}}} + \underbrace{\frac{\hat{k}}{c} \bullet \left(\dot{V}_t - \dot{V}_r\right)}_{\substack{\text{NORMAL} \\ \text{DOPPLER}}} + \underbrace{\ldots}_{\substack{\text{HIGHER ORDER} \\ \text{TERMS } (V/c)^3\text{, etc.}}} = 1 + \delta$$

- OBSERVED SATELLITE INCREASE IN FREQUENCY (OBSERVED AT RECEIVER)

  $\delta = 447.9 \times 10^{-12}$ or 448 μsec/sec OF TIME OFFSET (38.7 μsec/day)

- SOLAR PERTURBATION: VARIES FROM 445.8 to 450.2×10^{-12}

- SATELLITE CLOCK IS PURPOSELY SET LOW TO 10.22999999545 MHz OR 4.45×10^{-10} LOW IN FREQUENCY

1.6 Ionospheric and Tropospheric Range Errors

Both the ionosphere and the troposphere generate range errors. The ionospheric error is caused by the integrated electron count over the ray path after one has accounted for the ray bending effects of the ionosphere. Thus the effect is dependent on both the character of the ionosphere at zenith and the elevation angle to the satellite.

Figure 1-10 illustrates the typical effects of the elevation angle relative to the total ionospheric delay. The obliquity factor gives the factor with which the ionospheric delay is increased relative to the delay for a ray to a satellite at zenith. As shown, the obliquity factor is on the order of 3 for low elevation angles.

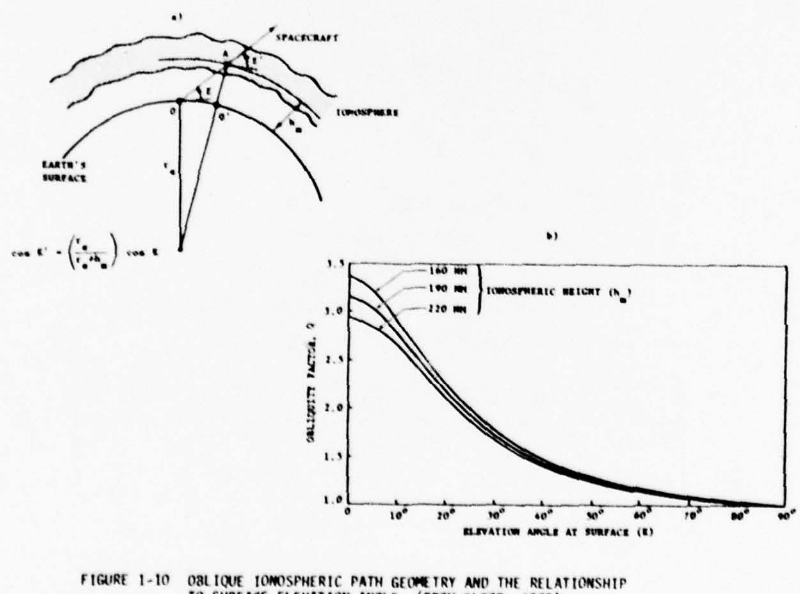


FIGURE 1-10 OBLIQUE IONOSPHERIC PATH GEOMETRY AND THE RELATIONSHIP TO SURFACE ELEVATION ANGLE, (FROM ELROD, 1975)

Figure 1-11 shows typical measurements of ionospheric delay for an L-band signal received at vertical incidence (obliquity factor of 1). The mean ionospheric delay at nighttime is on the order of 10 nsec. During daytime the delay increases to as high as 50 nsec. In regions near the geomagnetic equator or near the poles the delays can be significantly larger, particularly during periods of magnetic storms.

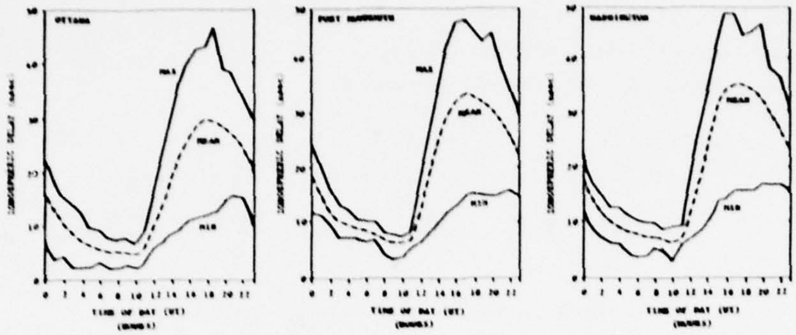


FIGURE 1-11 MEAN IONOSPHERIC DELAY AND ENVELOPE OF DELAY VARIATION vs. TIME OF DAY DURING - MARCH, 1958 - SATELLITE AT ZENITH f = 1.6 GHz (FROM ELROD, 1975)

At low elevation angles when the ionosphere delay can be 3 times the values given above and thus can be on the order of 30 nsec at nighttime or 150 nsec during daytime. Some of this delay can be eliminated by ionospheric modeling. However, more precise corrections can be made by making use of measurements at two L-band frequencies, 1575.42 MHz = L1 and 1227.6 MHz = L2, and making use of the approximate inverse square-law behavior of the ionospheric group delay as shown in Table 1-2.

Table 1-2 IONOSPHERIC GROUP DELAY VARIATION WITH FREQUENCY

$$\tau_{GD} = \frac{R}{c} + \frac{A}{f^2} + \frac{B}{f^3} + \frac{C}{f^4} \cdots$$

$$\approx \frac{R}{c} + \frac{A}{f^2}$$

where R = TRUE RANGE

C = SPEED OF LIGHT

f = CARRIER FREQUENCY

B= (AVERAGED EARTH MAGNETIC FIELD STRENGTH ALONG THE PATH)[3]

TERM NEGLECTED	RANGE ERROR AT f = 1.5 GHz
B/f^3	≤ 1 inch
C/f^4	≤ 3 inches

The excess delay caused by the ionosphere is mainly contributed by the term A/f^2. The B/f^3 term caused by magnetic field effects on the ionosphere and C/f^4 term generated by the binominal expansion are both small in significance. Thus one measures the total delay at both L1, L2 and computes the difference $\Delta\tau$ between the total group delays at F_{L1} and F_{L2}. The only difference between these two total group delays is caused by the frequency dependent ionospheric group delays, τ_{GDL1}, τ_{GDL2}. Thus the difference $\Delta\tau$ in total propagation delays reduces to

$$\Delta\tau \triangleq \tau_{GDL2} - \tau_{GDL1} = A\left[\frac{1}{f_{L2}^{2}} - \frac{1}{f_{L1}^{2}}\right] = \frac{A}{f_{L1}^{2}}\left[\frac{f_{L1}^{2} - f_{L2}^{2}}{f_{L2}^{2}}\right]$$

or

$$\Delta\tau \triangleq \tau_{GDL2} - \tau_{GDL1} = \tau_{GDL1} \left| \left(\frac{f_{L1}}{f_{L2}} \right)^2 - 1 \right| \qquad (1\text{-}5)$$

Thus a computation of $\Delta\tau$ gives an estimate of τ_{GDL1}.

The other significant contributor to delay error is the ray bending effects of the troposphere caused by water vapor and other atmospheric constituents. Figure 1-12 shows the frequency independent tropospheric effect and its sensitivity to the elevation angle and user altitude. This effect can also be significant at low elevation angles. However, the tropospheric effect is more easily predicted by making use of relatively simple atmospheric measurements at sea level and should not contribute a large residual error.

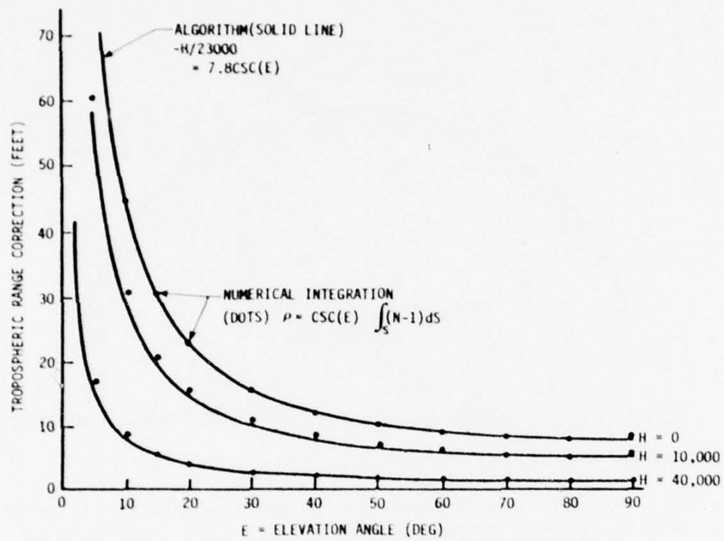


FIGURE 1-12 TROPO RANGE CORRECTION (COMPARISON OF ALGORITHM WITH NUMERICAL INTEGRATION) VS USER ALTITUDE H IN FEET

1.7 Multipath and Interference Effects

Many of the navigation users are in aircraft and often are above water where a substantial multipath reflection exists. See Fig. 1-13. When the satellite is at zenith the differential path delay is equal to twice the aircraft altitude, perhaps 1000 to 80,000 ft, or 2 to 160 μsec. At lower elevation angles, the differential delay might be 1/10 of this value. The magnitude of the reflected ray can sometimes be almost as large as the direct ray. The GPS signal should be tolerant to this multipath reflection.

Spurious interference from harmonics of narrow band CW transmitters is another possible interference source. The signal from the satellite received by an omni-directional antenna (0 dBIC) is at -130 dBm for the C/A signal. Many potential navigation signals cannot tolerate an interference level greater than 1/10 the received signal strength and certainly not one equal to the received signal power. Some signals, on the other hand, can tolerate much larger interference power.

The power received from an isotropic interference transmitter to our isotropic receive antenna is

$$P_r = P_t/(4.56 \times 10^3 f^2 d^2) \qquad (1\text{-}6)$$

where f is in MHz and d is in miles. For $f = 1.57542 \times 10^3$ MHz, $d = 10^2$ miles, $P_t = 10$ dBm. We have $P_r = P_t/1.13 \times 10^{14} = -140.5$ dB + P_t = -130.5 dBm. Hence even a 10 mw transmitter, 100 miles away, could interfere with the received satellite signal if it were not designed properly. Thus tolerance to low level in-band interference is an important aspect in GPS signal selection.

1.8 Summary of Desired Navigation Signal Characteristics

In the preceding paragraphs we have attempted to examine in turn each of the desired characteristics of the GPS navigation signal. This signal must be an RF representation of the satellite clock and in addition must carry data to indicate satellite position and clock correction parameters.

We summarize these requirements below and in Table 1-3.

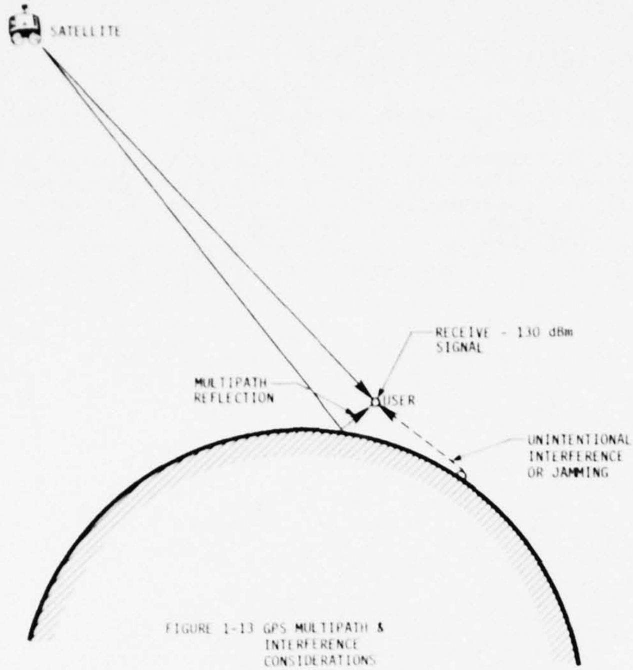


FIGURE 1-13 GPS MULTIPATH & INTERFERENCE CONSIDERATIONS

- Allow accurate pseudo-range measurements (σ_t<10 nsec) without ambiguity.
- Allow accurate doppler shift measurements (<0.1 Hz).
- Provide dual frequency measurements to provide ionospheric group delay measurements (> 20% frequency separation).
- Provide an efficient data channel for the transmission of satellite ephemeris, clock correction information, and other data (50 bps).
- Provide both a high accuracy "protected" signal along with a simpler signal which provides somewhat lower accuracy and is easily acquired in a short time (≈1-2 minutes) by the receiver.
- Good multiple access properties. The user will typically receive simultaneous transmissions from 6-11 satellites.
- Ability to resist interference from low power narrow band interference as well as moderate power intentional interference.
- Ability to reject or to reduce greatly multipath interference problems where the differential multipath delay is 200 nsec or greater.

Table 1-3 DESIRED GPS SIGNAL PROPERTIES

- ALLOW ACCURATE REAL-TIME TIME-OF-ARRIVAL MEASUREMENT (σ_t<10 nsec) WITHOUT AMBIGUITY.
- ALLOW ACCURATE DOPPLER SHIFT MEASUREMENT.
- PROVIDE AN EFFICIENT DATA CHANNEL.
- PROVIDE A RAPID ACQUISITION NAVIGATION CAPABILITY WITH GOOD ACCURACY ALONG WITH A HIGH ACCURACY CAPABILITY FOR MORE DEMANDING USERS.
- PROVIDE IONOSPHERIC GROUP DELAY CORRECTION.
- GOOD MULTIPLE ACCESS PROPERTIES.
- GOOD INTERFERENCE REJECTION PROPERTIES.
- TOLERANCE TO MULTIPATH INTERFERENCE.
- SIGNAL GENERATION COMPATIBLE WITH CURRENT SPACE ELECTRONICS TECHNOLOGY.
- AVOID EXCESSIVE BANDWIDTH RELATIVE TO THE CENTER FREQUENCY.

SECTION 2

GPS SIGNAL STRUCTURE

2.1 Introduction

In this section the structure of the GPS signal is described in detail. The general properties of the codes employed are discussed along with a consideration of many of the system requirements, e.g., multiple access discussed in Section 1. The remaining requirements are discussed later in Section 3.

2.2 GPS Signal Frequency Characteristics

The GPS signal consists of two components, Link 1, L1, at a center frequency of 1575.42 MHz and Link 2, L2, at a center frequency of 1227.6 MHz. The L-band center frequency selection has advantages over lower frequencies in that the channel bandwidth allocation is more readily obtainable at L-band. In addition, the ionospheric delay effects (without correction) are substantially smaller. As compared to C-band, the space losses to an isotropic receive antenna are substantially larger for C-band than L-band and give L-band the advantage. Thus the frequency separation is 347.82 MHz or 28.3% relative to L2. As discussed earlier, this dual frequency measurement permits measurement of the ionospheric group delay error. Each of these center frequencies is a coherently selected multiple of a 10.23 MHz clock. The frequency stability for the NTS-2 clock with redundant cesium standards is better than 2 parts in 10^{13}. In particular the link frequencies are:

$$\begin{aligned} L1 &= 1575.42 \text{ MHz} = 154 \times 10.23 \text{ MHz} \\ L2 &= 1227.6 \text{ MHz} = 120 \times 10.23 \text{ MHz} \end{aligned} \qquad (2\text{-}1)$$

Thus the ionospheric group delay correction equation (1-5) of Section 1 becomes

$$\Delta\tau \triangleq \tau_{GDL2} - \tau_{GDL1} = \frac{A}{f_{L1}^{\,2}} \frac{1}{1.5336} = \frac{\tau_{GDL1}}{1.5336} \qquad (2\text{-}2)$$

or

$$\tau_{GDL1} = 1.5336\ \Delta\tau$$

where τ_{GDL1} is the ionospheric group delay at L1 and $\Delta\tau$ is the measurable difference between total propagation delays at L1, and L2.

As discussed earlier, the relativistic effects are partially compensated for in the satellite by offsetting the 10.23 MHz clock slightly low by a factor of 4.45×10^{-10} or 4.55×10^{-3} Hz at the 10.23 MHz clock rate. Thus, as the signal approaches the earth from the satellite, the frequency increases by approximately the same factor and the signal appears to a stationary user on the earth to have a frequency very close to 10.23 MHz. Henceforth when reference is made to 10.23 MHz, the frequency will always be this offset frequency as far as satellite clocks are concerned.

Each of these two signals L1 and L2, is modulated by either or both a 10.23 MHz clock rate precision (P) signal or by a 1.023 MHz clear/acquisition (C/A) signal. Each of these two binary signals has been formed by a P-code or a C/A code which is modulo-2 added to 50 bps data D, to form P⊕D and C/A⊕D, respectively.

The L1 in-phase component of the carrier is modulated by the P signal P⊕D and the quadrature carrier component is modulated by C/A⊕D. A phasor diagram of the L1 signal is shown in Fig. 2-1.

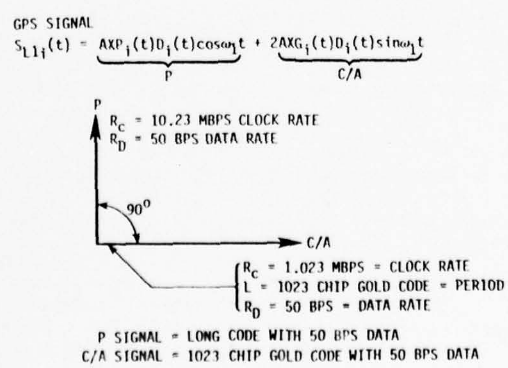


FIGURE 2-1 GPS SIGNAL STRUCTURE FOR L1 SIGNAL

The L2 signal is biphase modulated by either the P code or the C/A code. Normal operation would provide P code modulation on the L2 signal. The transmitted signal spectrum showing both L1 and L2 is shown in Fig. 2-2.

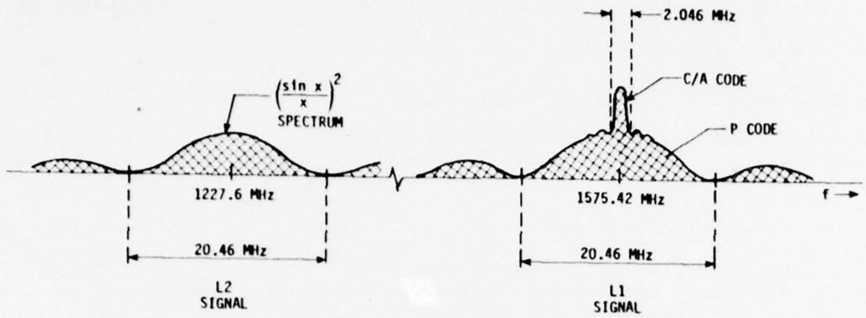


FIGURE 2-2 GPS SIGNAL POWER SPECTRAL DENSITY

2.3 Detailed Signal Structure

The Link 1 signal L1 contains both in-phase and quadrature signals. The signal transmitted (see Fig. 2-1) by the satellite is then

$$S_{L1i}(t) = A_P \, XP_i(t) \, D_i(t) \cos(\omega_1 t+\emptyset) + A_c \, XG_i(t) D_i(t) \sin(\omega_1 t+\emptyset) \qquad (2\text{-}3)$$

where ω_1 is the L1 frequency as defined above, $\emptyset$ represents a small phase noise and oscillator drift component. Oscillator stability is obtained using redundant cesium or rubidium frequency standards. (The first satellite in the GPS series NTS-2, has a clock stability better than 2×10^{-13}). The P-code, $XP_i(t)$, is a ± 1 pseudo-random sequence with a clock rate of 10.23 Mbps and a constrained period of exactly 1 week. Each satellite, i, transmits a unique P-code. The data, $D_i(t)$, also has amplitude ± 1 at 50 bps and has a 6 sec subframe and a 30 sec frame period. The C/A code XG_i is a unique Gold code of period 1023 bits and has a clock rate of 1.023 Mbps. Thus the C/A code has a period of 1 msec. The relative amplitudes of the P and C/A codes are controlled by the constants A_P and A_C. In GPS Phase 1 the C/A code strength is between 3 and 6 dB stronger than the P-code. As already mentioned above, the code clocks and transmitted RF frequencies are all coherently derived from the same on-board satellite frequency standard. The rms clock transition time difference between the C/A and P-code clocks is less than 5 nsec.

P-Code

The P-code for each satellite i is the product of 2 PN codes, X1(t) and $X2(t+n_i t)$, where X1 has a period of 1.5 sec or 15,345,000 chips and X2 has a period of 15,345,037 chips or 37 chips longer; both sequences are reset to begin the week at the same epoch time. Both X1 and X2 are clocked in phase at a rate 1/T = 10.23 MHz. Thus, the P-code is a product code of the form

$$XP_i(t) = X1(t) \, X2(t+n_i T) \text{ , Reset at beginning of week.} \qquad (2\text{-}4)$$

where the delay between X1(t) and X2(t) is n_i code clock intervals of T sec each. Each satellite has a unique code offset $n_i T$ which makes the P-code unique as well. The increase in code period for X2 by 37 relative to X1 allows the values of n_i to range over 0 to 36 without having any significant segment of a P-code of one satellite match that of another. Thus we have 37 different P-codes.

The period of a product of P-codes each of relatively prime period is the product of the periods. Thus if the P-code were allowed to continue without being reset it would continue without repetition for slightly more than 38 weeks. This overall period has been in effect subdivided so that each satellite gets a one week period which is non-overlapping with every other satellite.

The Z-count is defined as the number of 1.5 sec X1 epochs since the beginning of the week. Thus there are 4 X1 epochs per data subframe of 6 sec. In order to acquire the P-code, the 50 bps data stream contains a new Hand-Over-Word (HOW) each 6 sec subframe. The HOW word, when multiplied by 4, gives the Z-count at the beginning of the next 6 sec subframe. Thus if one knows the subframe epoch times and the HOW word, one can acquire the P-code at the next subframe epoch.

Figure 2-3 summarizes the timing relationships between X1, X2 epochs and the Z-count and HOW words.

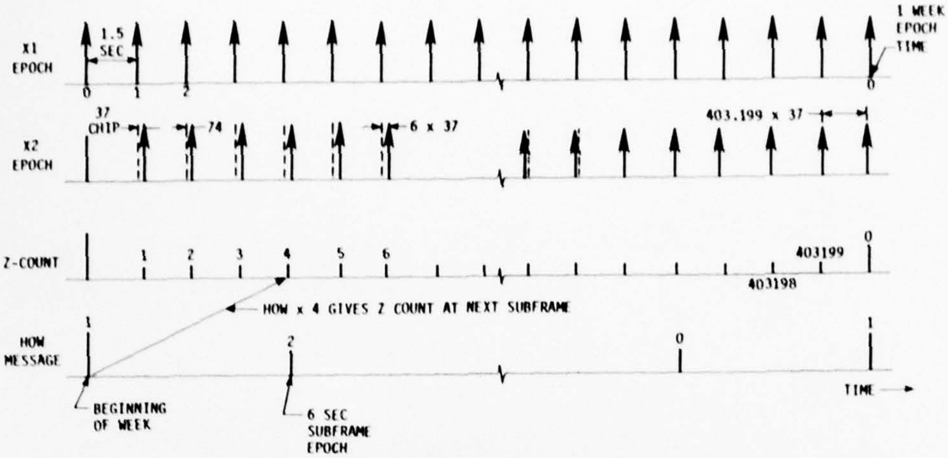


FIGURE 2-3 TIMING DIAGRAM FOR THE P CODE COMPONENTS X1, X2, AND THE Z-COUNT AND HOW MESSAGE RELATIONSHIP. THE HOW MESSAGE IS CARRIED IN THE 50 bps DATA STREAM

C/A Code

The clear/acquisition C/A code is a relatively short code of 1023 bits or 1 msec duration at a 1.023 Mbps bit rate. This code is selected to provide good multiple access properties for its period. The C/A codes for the various satellite are formed as the product of two 1023 bit PN codes G1(t) and G2(t). Thus this product code is also of 1023 bit period and is represented as

$$XG(t) = G1(t)\ G2\left|t + N_i(10T)\right| \qquad (2\text{-}5)$$

where N_i determines the phase offset in chips between G1 and G2. Note that a C/A code chip has duration 10T sec. There are 1023 different offsets N_i and hence 1023 different codes of this form.* Each code G1, G2 is generated by a maximal-length linear shift register of 10 stages. The G1 and G2 shift register, are set to the all ones state in synchronism with the X1 epoch. The tap positions are specified by the generator polynomial for the two codes

$$G1: \quad G_1\ (X) = 1 + X^3 + X^{10}$$

$$G2: \quad G_2\ (X) = 1 + X^2 + X^3 + X^6 + X^8 + X^9 + X^{10} \qquad (2\text{-}6)$$

Since the Gold code has a 1 msec period, there are 20 C/A code epochs for every data bit. The 50 bps data clock is synchronous with both the C/A epochs and the X1 epochs.

Figure 2-4 shows a simplified block diagrams of the C/A code generator. The unit is comprised of two 10 stage feedback shift registers clocked at 1.023 Mbps having feedback taps at stages 3, 10 for G1 and at 2,3,6,8,9,10 for G2. The various delay offsets are generated by tapping off at approximate points on the G2 register and modulo-2 adding the two sequences together to get the desired delayed version of the G2 sequence.

Epochs of the G code at 1 Kbps are divided down by 20 to get the 50 bps data clock. All clocks are in phase synchronism with the X1 clock as shown in Fig. 2-4.

L2 Signal

The L2 signal is biphase modulated by either the P-code or the C/A code as selected by ground command. The same 50 bps data stream modulates the L2 carrier as is transmitted on L1. Thus the L2 signal is represented in the normal P format as

$$S_{L2}(t) = B_p\ XP_i(t)\ D_i(t)\ \cos(\omega_2 t + 4) \qquad (2\text{-}7)$$

where B_p represents the signal amplitude at the satellite, $XP_i(t)$ is the P-code for the ith satellite clocked in synchronism with the L1 codes. Both carrier and code are synchronous with one another.

*There actually are 1025 different Gold codes of this period and family. The codes, G1(t) and G2(t), by themselves, are the other two codes.

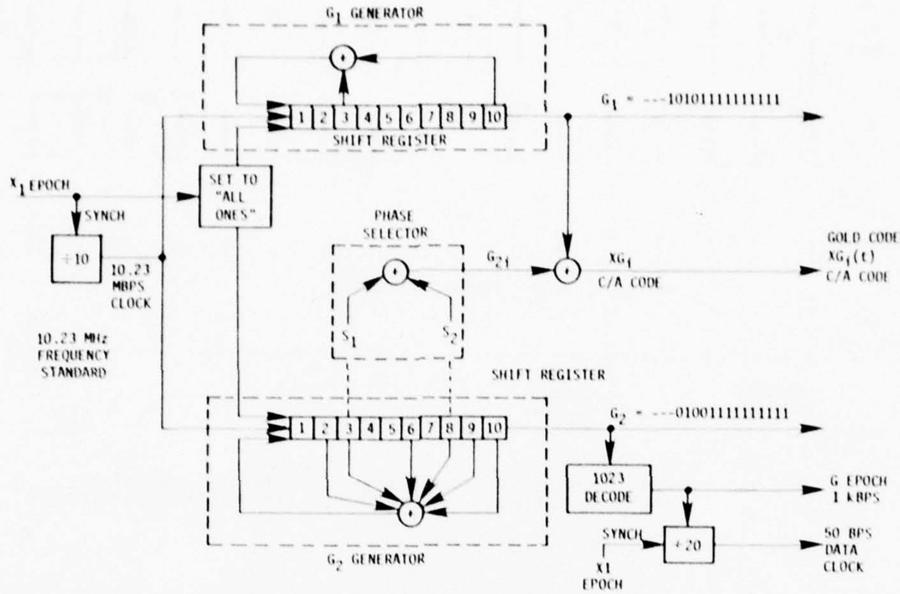


FIGURE 2-4 C/A CODE GENERATION

GPS Signal Summary

Table 2-1 summarizes the signal characteristics discussed above. One of the key points to be made in the signal structure discussion is that acquisition by a receiver of the relatively short period C/A code and recovery of a single full subframe of data permits one to acquire the P-code with minimal or zero search. Knowledge of the C/A epoch plus the data subframe epoch and the HOW word gives the exact phasing of the P-code.

Table 2-1 SUMMARY OF GPS SIGNAL PARAMETERS

PARAMETER	C/A SIGNAL	P SIGNAL
CODE CLOCK RATE, R_c	1.023 Mbps	10.23 Mbps
CODE LENGTH	1023	~6 X 10^{12}
DATA RATE, R_D	50 bps	50 bps
TRANSMISSION FREQUENCY	L1	L1, L2
DATA INCLUDES:	TELEMETRY SATELLITE EPHEMERIS SATELLITE CLOCK CORRECTION IONOSPHERIC MODEL P-SIGNAL ACQUISITION WORD ALMANAC	

The received signal strength at a user receiver employing a 0 dBIC antenna is given below in Table 2-2 for the GPS Phase I.

Table 2-2 GPS Received Signal power levels at output of a 0 dBIC antenna with RH circular polarization. The satellite is at an elevation angle $\geq 5^{\circ}$.

	GPS Signal Component (Minimum Strength)	
Link	P	C/A
L1	-163 dBw	-160 dBw
L2	-166 dBw	-166 dBw

The signal power spectral densities for the P and C/A signal components are shown in Fig. 2-5. Figure 2-6 shows the measured RF power spectral density of the L1 signal. Note the narrow band high power density C/A signal in the center of the signal spectrum.

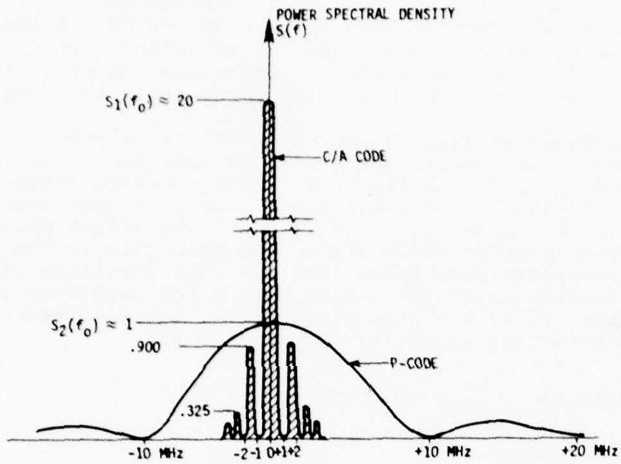


FIGURE 2-5 SPECTRA OF CARRIERS WITH BIT RATES OF 1 MEGABIT/SEC AND 10 MEGABITS/SEC. THE RATIO OF C/A POWER TO P-CODE SIGNAL POWER IS 3 dB IN THIS FIGURE.

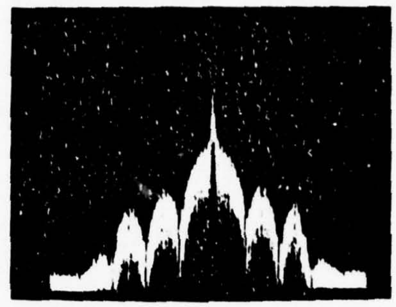


FIGURE 2-6 RF SPECTRUM OF RECEIVED L1 CARRIER WITH C/A AND P QSPK MODULATION, GENERATED BY STI TEST TRANSMITTER

SPECTRUM SCALES:
HORIZONTAL: 10 MHz/DIVISION
VERTICAL: 10 dB/DIVISION

2.4 Signal Characteristics

In the previous paragraph we defined the structure of the GPS signal but did not more than begin to examine its characteristics and performance. In this section we begin to examine the multiple access characteristics of the C/A and P-codes and to give some of the reasons for their selection. In particular, the cross-correlation properties are examined for the P and C/A signals both with and without doppler offset.

Cross-Correlation Properties

The key multiple access performance parameter of the GPS signals is the generalized cross-correlation performance. Any GPS receiver must in effect perform a cross-correlation operation if it is to extract the signal and recover the data.

Figure 2-7 shows the typical received signal format and cross-correlation receiver where two satellites are in view, satellite h, the desired signal, and satellite j, the interfering satellite. Of course, one must realize that a parallel or time multiplexed correlator will next be reversing these roles and satellite j would be the desired signal and satellite h would be the interference. In general, of course, more than 2 satellites are in view and the receiver operates on at least 4 satellite signals.

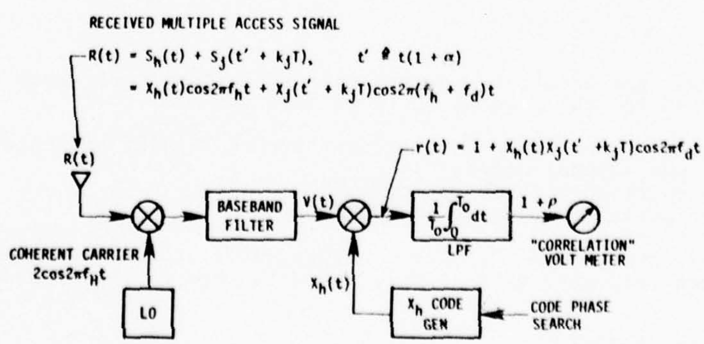
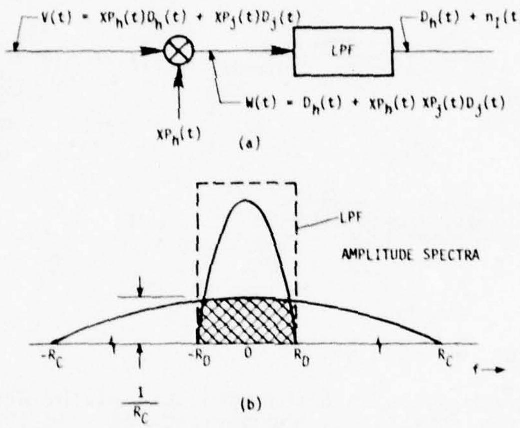


FIGURE 2-7 MULTIPLE ACCESS INTERFERENCE IN USER RECEIVER - THE RECEIVED SIGNAL CONSISTS OF THE DESIRED SIGNAL PLUS A TIME OFFSET, DOPPLER SHIFTED MULTIPLE ACCESS SIGNAL FROM ANOTHER SATELLITE.

The received signals are assumed to be modulated by codes $X_h(t)$ and $X_j(t) = \pm 1$ respectively. These signals can represent either the P-code or the C/A code or an arbitrary signal. For the moment the data modulation is ignored and the signals are of equal strength. Noise effects are additive and can be considered separately.

The block diagram of Fig. 2-7 shows a coherent correlation operation where the coherent carrier is multiplied with the received signal and the resultant baseband output is multiplied by a phase synchronized replica of the desired code $X_h(t)$. The output of the multiplier is then integrated for some time T_m sec to produce the "correlation" output. If T_m is equal to or a multiple of the period of the waveforms or approaches infinity, the output will be the true correlation, otherwise it will be a partial correlation function.*

The output of the "correlation" meter would be exactly zero if there is no cross-correlation between X_h and X_j codes. However, in general there will be some finite cross-correlation either positive or negative and $1 + P \neq 1$ in Fig. 2-8. This non-zero cross-correlation can cause interference in the receiver or possible false lock in a code search and acquisition operation if $|P|$ is sufficiently large, e.g., $|P| > 0.3$. The effect can be made more severe by the user receiver antenna pattern which might have more gain in the direction of the interfering satellite and perhaps less space loss for that satellite as well. For example, if the interfering satellite is at the zenith and in the direction of maximum antenna gain while the desired signal is at a 5° elevation angle, the difference in received signal levels can favor the interfering signal by more than 6 dB.



SIGNAL-TO-INTERFERENCE RATIO

$$\frac{S}{I} = \frac{P_S(h)}{\frac{P_S(j)}{R_C} 2R_D} = \frac{R_C}{2R_D}$$

$\frac{S}{I} = 50$ dB FOR THE GPS P - SIGNAL AND $P_S(h) = P_S(j)$

(c)

FIGURE 2-8 MULTIPLE ACCESS GAIN

P-Code

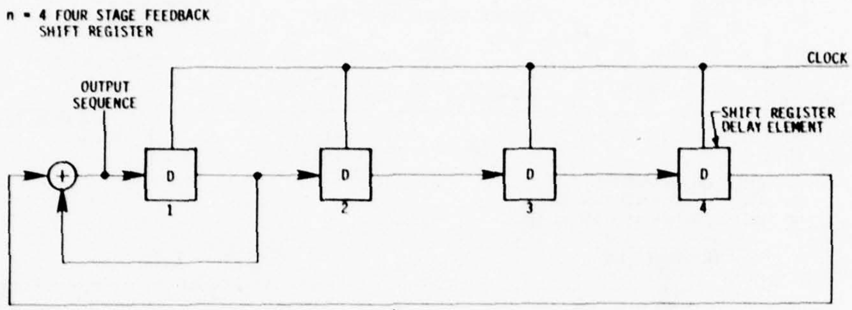
Figure 2-8 shows the amplitude spectra and signal-to-interference ratio (multiple access gain) computed for the P-codes where we have assumed:

- The desired and multiple access interference signals are received at equal power and the same doppler offset
- Both are received with 50 bps data
- The two signals are clocked in synchronism

The output spectrum at the multiplier output, $W(t)$, then takes the form shown in Fig. 2-8(b). The desired component gives simply the data spectrum $D_h(t)$ here assumed to be a

*Spilker, op cit, pp. 597-600

random data stream. The interfering multiple-access component has the spectrum of the product of the two PN codes which is for all purposes of interest here, a pseudo-random bit stream. Thus both have $(\sin x/x)^2$ spectra, one with bandwidth to the null of R_d = 50 Hz and the other at 10.23 MHz. Clearly one can place a low pass filter of bandwidth R_d and get most of the power (the added power by using a larger bandwidth would be less than 0.44 dB). The power passing to the output of the low-pass filter from the multiple access interference is reduced by a factor of $2R_d/R_c$ where R_c = 10.23 MHz. Thus the output signal-to-interference power ratio for the P-code is $S/I = +10^5$ or 50 dB.



(a)

SEQUENCE OF STATES TIME →

STATE 15 = $2^n - 1$

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16		
1	0	0	0	1	1	1	1	0	1	0	1	1	0	0	1	0	0	0
2	1	0	0	0	1	1	1	1	0	1	0	1	1	0	0	1	0	0
3	0	1	0	0	0	1	1	1	1	0	1	0	1	1	0	0	1	0
4	0	0	1	0	0	0	1	1	1	1	0	1	0	1	1	0	0	1

4 BIT STATES

(b)

FIGURE 2-9 GENERATION OF A PN SEQUENCE

2.5 C/A Code Properties

The multiple access properties of the Gold codes are substantially different from the P-code signal in two respects:

- Cross-correlation sidelobes are not of equal height and are much larger than those of the P-code.
- The cross-correlation property is dependent in a significant way on both doppler offset as well as code offset. The P-code on the other hand has multiple access properties essentially independent of doppler and time offset.

Linear Feedback Shift Register Sequence

Before we discuss details of the Gold codes, we review briefly the properties of the PN codes. Figure 2-9 shows a 4-stage linear maximal length shift register and the sequence it generates at each clock pulse. The sequence of shift register states is shown in Fig. 2-9 (b), the initial state is

$$S = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \tag{2-9}$$

where the state vector components are defined as the state of each of the binary shift register delay elements. As long as this shift register is not forced to the "all zero" state, it will cycle through all $2^4-1 = 15$ states in a periodic manner. In general, an n stage linear feedback shift register (LFSR) with proper taps produces a code of period $P = 2^n-1$. Note that the LFSR generates all state factors except the all zero state, thus it cycles through each of the possible state vectors. Thus there are 2^n-1 states in the period.

Figure 2-10 shows the PN sequence at the output of a selected stage. The autocorrelation of the PN sequence where $s(t) = \pm 1$ is

$$R(i) = (1/15) \int_0^{15} s(t)\, s(t+i)\, dt \tag{2-10}$$

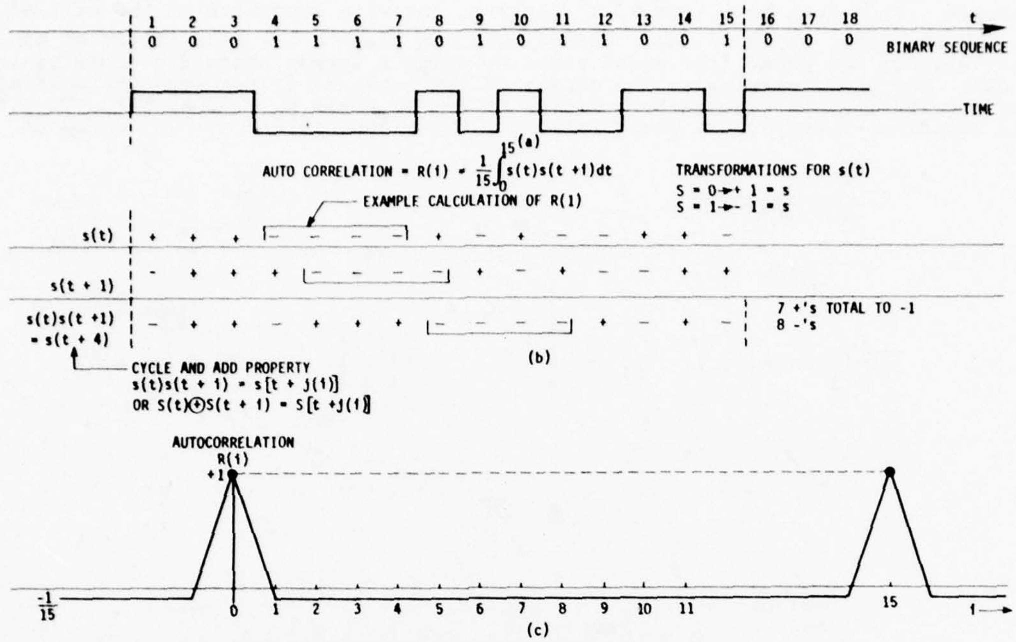


FIGURE 2-10 AUTOCORRELATION FUNCTION OF A PN SEQUENCE

For a unit time offset i = 1, the product s(t) s(t + i) is shown in Fig. 2-10(b). This product is easily seen to form a shifted version of the same sequence. This property is variously known as the shift-and-add or cycle-and-add property in reference to the fact that shifting the sequence S(t) = (0,1) by i clock pulses and modulo-2 addition with the original unshifted sequence forms a shifted version of the same sequence with a different offset.

Since the PN sequence has one more -1 than +1, the average value of s(i) is simply $-1/(2^n-1) = -1/15$ where $2^n -1$ is the code period. Thus, it is easily seen that the autocorrelation function of a LFSR is a two level function as shown in Fig. 2-10(c) for integral values of i.

$$R(i) = 1 \qquad i = 0$$
$$= \frac{-1}{2^n-1} \qquad i \neq 0 \tag{2-11}$$

For the autocorrelation function for nonintegral values of time offset, simply connect the values of R(i) by straight lines.

Gold Codes

The Gold codes* are a family of codes formed as the product of two different LFSR, both of the same period $P \triangleq 2^n-1$. Table 2-3 illustrates some of the properties of a Gold code. The two PN sequences $X_i(t)$, $X_j(t)$ are specially selected from the set of LFSR sequences having the same period.

Using Table 2-3 it is easily seen that the cross-correlation between any two different Gold codes of the same family $G_k(t)$ and $G_e(t)$ with no time offset is simply -1/P the same as for the PN code autocorrelation. More generally, however, the cross-correlation

$$\overline{G_k(t)\ G_e(t+n)} = \overline{G_s(t+r)} = \overline{G_s(t)} \tag{2-12}$$

is simply the time average of another code in the same family. Table 2-4 summarizes the quantitative results for cross-correlation with zero doppler offset.

Table 2-5 summarizes the performance of three types of sequences; linear maximal length shift register sequences, nonlinear maximal length shift registers (contains one more state in the period since they have all zero states) and the Gold codes. Note that there are 2^n+1 Gold codes in a family of codes of period 2^n-1. There are all shift offsets k allowed as indicated in Table 2-3 plus the two PN components by themselves $X_i(t)$, $X_g(t)$.

*R. Gold, "Optimal Binary Sequences for Spread Spectrum Multiplexing," IEEE Trans. on Info. Theory, Oct. 1967, pp. 619-621.

TABLE 2-3 GOLD CODE PROPERTIES

PRODUCT CODES - PRODUCT OF 2 PN CODES OF SAME PERIOD

$G_k(t) = x_i(t)\ x_j(t+k)$ WHERE x_i HAVE PERIOD P

P DIFFERENT VALUES OF k. HENCE P DIFFERENT CODES PLUS x_i, x_j

FAMILY OF CODES GENERATED FOR DIFFERENT VALUES OF k WITH LOW CROSS-CORRELATION

CROSS CORRELATION $\overline{G_k(t)G_\ell(t)} = \overline{x_i(t)x_j(t+k)\alpha_i(t)x_j(t+\ell)} = \overline{x_j(t+k)x_j(t+j)}$

$$= \overline{x_j(t+m)} = \frac{-1}{P}$$

USE CYCLE AND ADD PROPERTY OF PN SEQUENCE

FOR OTHER VALUES OF CODE SHIFT $G_k(t)G_\ell(t+n)$ THE CROSS CORRELATION IS BOUNDED

$$\overline{G_k(t)G_\ell(t+n)} = \overline{x_i(t)x_j(t+k)x_i(t+n)x_j(t+k+n)}$$

$$= \overline{x_i(t+r)x_j(t+r+s)} = \overline{G_s(t+r)} = \overline{G_s(t)}$$

TABLE 2-4 CROSS CORRELATION PROPERTIES OF GOLD CODES

CODE PERIOD	NUMBER OF SHIFT REGISTER STAGES	NORMALIZED CROSS-CORRELATION LEVEL	PROBABILITY OF LEVEL
$P = 2^n - 1$	n	$-\frac{\left(2^{\left(\frac{n+1}{2}\right)}+1\right)}{P}$	0.25
	n-ODD	$-\frac{1}{P}$	0.50
		$\frac{\left(2^{\left(\frac{n+2}{2}\right)}-1\right)}{P}$	0.25
$P = 2^n - 1$	n	$-\frac{\left(2^{\left(\frac{n+2}{2}\right)}+1\right)}{P}$	0.125
	n-EVEN	$-\frac{1}{P}$	0.75
	$n \neq 41$	$\frac{\left(2^{\left(\frac{2+n}{2}\right)}-1\right)}{P}$	0.125

Table 2-5 GENERALIZED PROPERTIES - MAXIMAL LENGTH CODES

	Linear Shift Registers	Nonlinear Shift Registers	Gold Codes
Period for n-Stages	2^n-1	2^n	$2^{n'}-1$ $n' = n/2$
Cycle and Add Property	Yes	No	Yes
Autocorrelation Function	Two-Level	Cannot be Two-Level	4 - Level
Number of Codes of Period	$N = \frac{\phi(2^n-1)}{n} \leq \frac{2^n-2}{n}$	$\approx \frac{2^{2^n}}{2^n}$	$2^{n'}+1$
2^n-1 or 2^n	For $2^n-1 = 8191$, N=630		8193

Where $\phi(m)$ is the number of integers relatively prime to m

Thus the advantage of the Gold codes is not simply a low cross-correlation between all members of the family but that there are a large number of codes all of similar good properties.

The amplitude spectral density of a PN code is shown in Figure 2-11. If the PN code period is 1023 bits and the clock rate is 1.023 Mbps, then the PN code spectrum is a set of line components with a (sin kf)/kf variation in amplitude level. The line components are, of course, separated by the inverse code period rate R_c/P = 1 kHz apart where R_c is the code clock rate and P is the period.

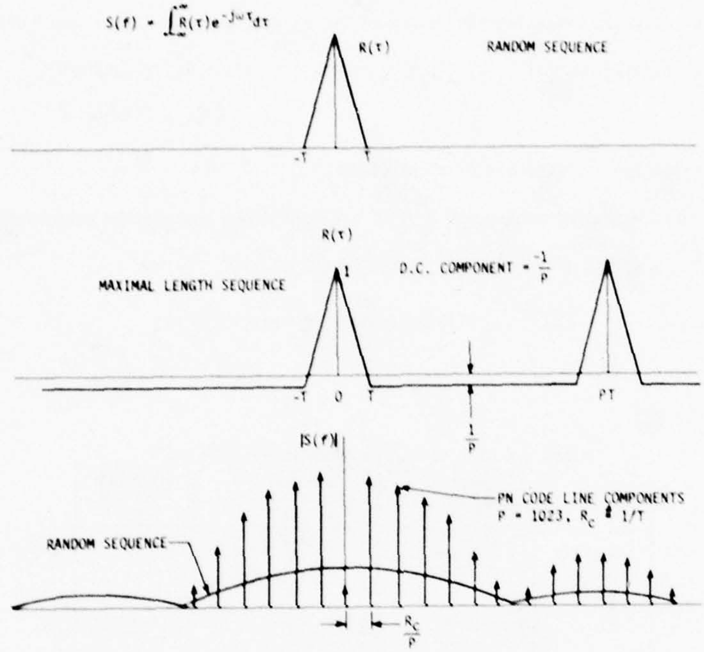


FIGURE 2-11 AMPLITUDE SPECTRUM OF SHORT PN CODE

The Gold codes of the same period are composed of a similar set of line components. However, in this instance, the line components are not all of the same amplitude although the frequency spacing is the same. Figure 2-12 shows an example spectrum for a code period P = 1023. Note that instead of a line component power 30 dB down from the total signal power, P_s, as would be if we had 1000 line components of equal power, the line component power level varies significantly about this level.

The cross-correlation between two Gold codes with both time offset and doppler offset is

$$\overline{G_k(t)\ G_e(t+n)\ \cos\ \omega_d t} = \overline{G_s(t)\ \cos\ \omega_d t}$$

where bar denotes the time average. Thus if the doppler offset is an integral multiple of the line component spacing, the cross-correlation is simply the amplitude of that line component. Recall that even with stationary use the doppler offset between satellites can vary by as much as 7500 Hz.

Table 2-6 summarizes the cross-correlation results for both zero doppler and the worst possible doppler. Note that the zero doppler cross-correlation changes by 6 dB at every 2 stage increase in the number of shift registers. Thus the zero doppler cross-correlation decreases by 6 dB by going from period P = 511 to P = 2047 whereas an increase from 511 to 1023 causes no improvement.

It is easily seen, however, that the zero doppler condition is not the one of greatest importance. When worst case doppler shifts are considered, the peak cross-correlation changes by 3 dB with each increase in code period.

Figure 2-13 shows the cumulative probability of various cross-correlation interference levels for the GPS C/A code for various doppler shifts from f_d = 0 to $\pm$ 5 kHz. Note that the 4 kHz doppler gives the worst cross-correlation sidelobe over this range; however, the other doppler shifts give similar results. These cumulative averages are formed by averaging results for all 1023 of the Gold codes of period 1023 in the GPS family. All possible code time offsets are considered for each doppler offset and all possible pairs of codes in this family.

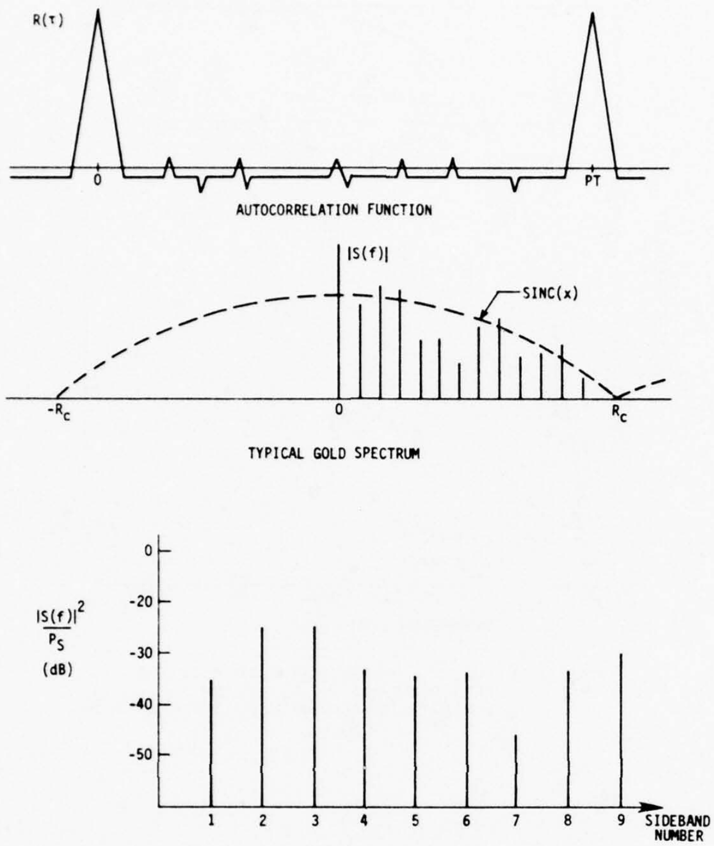


FIGURE 2-12 GOLD CODE SPECTRUM

Table 2-6 CROSS-CORRELATION SIDELOBES FOR GOLD CODES

Parameter	CODE PERIOD		
	511	1023	2047
Peak Cross-correlation (any doppler shift)	-18.6 dB	-21.6 dB	-24.6 dB
Peak Cross-correlation (zero doppler)	-23.8 dB	-23.8 dB	-29.8 dB
Probability of worst case or near worst case cross-correlation (zero doppler)	0.5	0.25	0.5

Note one additional point. If there is a doppler shift between two signals, then the delay difference is changing between codes. For example, if there is a 1 kHz doppler shift at L1, then the code C/A clock rate differs by 1 kHz/1540 = 1/1.54. Thus the two codes will shift in relative delay by one C/A code chip every 1.54 sec. Thus by their very nature these sidelobes with doppler shift are only temporary in nature.

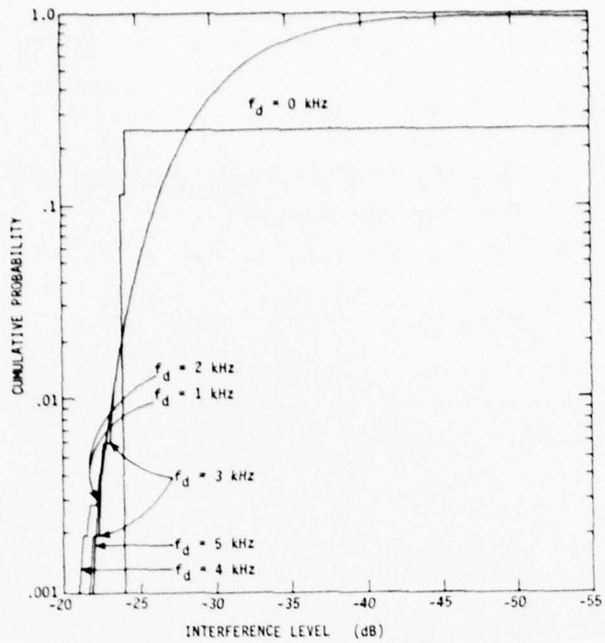


FIGURE 2-13 CUMULATIVE PROBABILITY OF INTERFERENCE LEVEL FOR 1023 bit GOLD CODE AT 1.023 Mbps (COURTESY H. CHANG, STANFORD TELECOMMUNICATIONS INC.)

SECTION 3

GPS SIGNAL TRACKING AND ACQUISITION

3.1 Introduction

Thus far, the GPS signal has been described and its tolerance to multiple access interference has been discussed. Described as well is its lack of time ambiguity and the dual frequency ionospheric correction capability.

Yet to be demonstrated, however, is the accuracy of the receiver tracking, the ease in acquiring the signal, the multipath interference and other interference rejection performance. In order to analyze this performance, it is first necessary to describe the basics of delay-lock loop receiver operation. There is no attempt here to analyze the performance of these receivers. These analyses are contained elsewhere already.* Rather, the attempt is to summarize the key performance results as applicable to the GPS signal.

3.2 Delay Lock Receivers

The receiver must accurately track the received GPS signals even though they are received at low signal levels, usually well below the thermal noise level in the receiver. In addition, the receiver must be able to track the dynamics of motion of the user platform, perhaps a high performance aircraft. As shown later, these two contending requirements can still be satisfied while producing the desired performance accuracy.

The essential elements of the delay-lock loop (DLL) is the correlator shown in Fig. 3-1. A received code is multiplied by a reference code time offset by $\tau < T$ where T is a code chip interval. The multiplier output V_3 is averaged by a low-pass filter having an integration time $T_m = 1/B \gg T$, i.e., much greater than the chip interval. For the moment, the code period is assumed to be essentially infinite. Thus, the output $V_4(\tau) = R_p(\tau)$ is the autocorrelation function of the code. The correlator output itself is not sufficient for code tracking, however, it does not provide an indication of the sign of the delay error of a tracking reference signal.

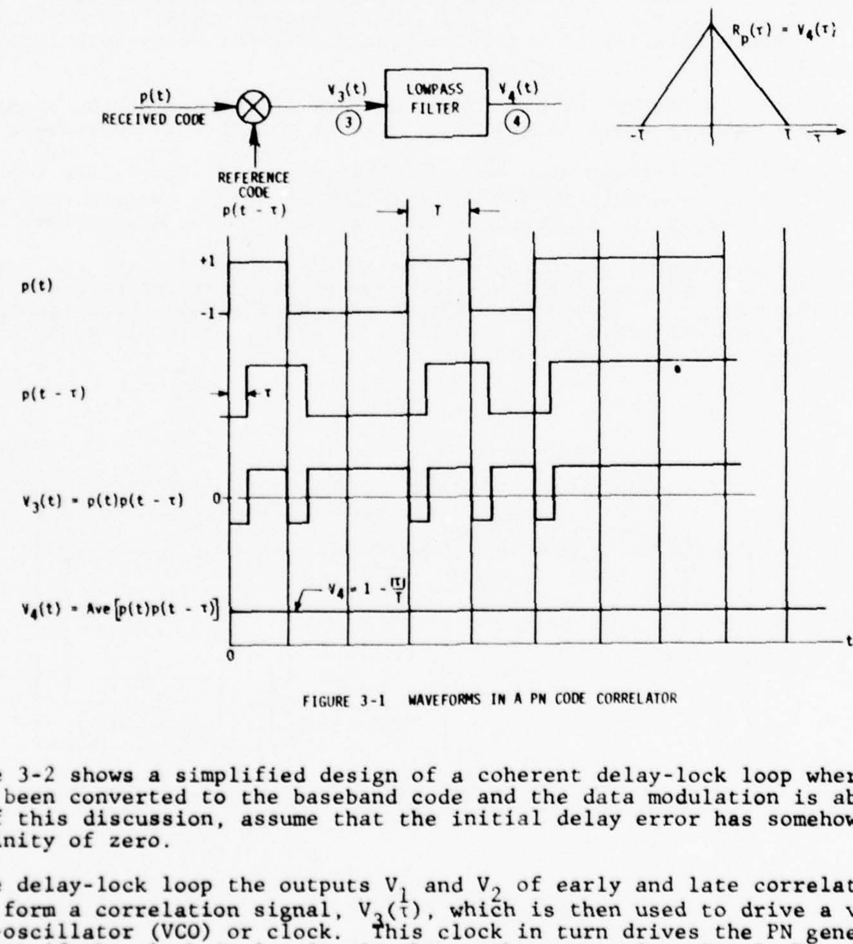


FIGURE 3-1 WAVEFORMS IN A PN CODE CORRELATOR

Figure 3-2 shows a simplified design of a coherent delay-lock loop where the received signal has been converted to the baseband code and the data modulation is absent. For purposes of this discussion, assume that the initial delay error has somehow been decreased to the vicinity of zero.

In the delay-lock loop the outputs V_1 and V_2 of early and late correlation are subtracted to form a correlation signal, $V_3(\tau)$, which is then used to drive a voltage-controlled-oscillator (VCO) or clock. This clock in turn drives the PN generator in such a manner that if the clock is lagging in phase, the correction signal, V_3, drives the clock faster and the reference code speeds up and runs in coincidence with the received signal. Thus the reference code is tracking the received code. The epoch time ticks are then a measure of the received signal time. The receiver also contains a coincident or punctual channel as is shown in the top portion of the block diagram in Fig. 3-2.

BASIC CODE TRACKING LOOP (DELAY - LOCK LOOP)

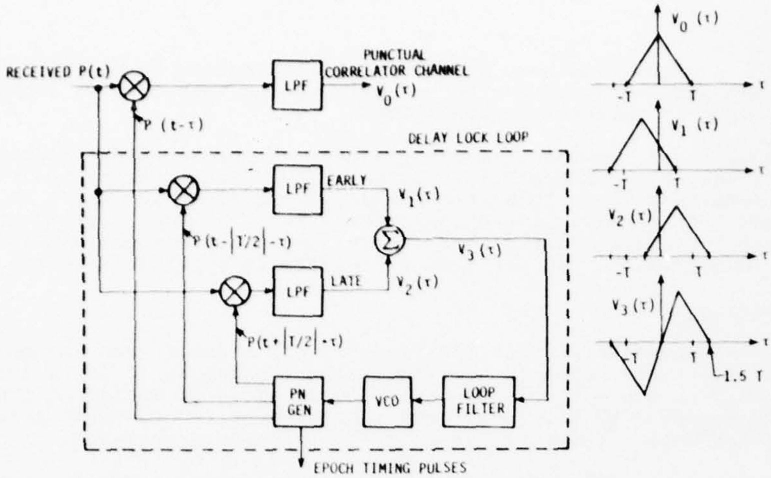


FIG.3-2 COHERENT DELAY LOCK LOOP

If the received signal delay increases suddenly because of user platform motion the delay error increases momentarily and the correction signal increases from zero. The reference code then slows down and increases its delay until it matches the received signal at which point the correction signal decreases to zero again. Thus it is clear that given an initial small error ($|\tau| < 1.5T$, the locked-on state) and sufficiently slow dynamics of delay change relative to the filter bandwidths the delay-lock loop will track the incoming signal.

If additive noise is present at the input the form of the correction signal of Fig. 3-2 does not change, however there is additive noise on top of this correction component.

If data modulation is present at a low rate the outputs V_1 and V_2 are replaced by $V_1D(t)$ and $V_2D(t)$. Since the data is ± 1 the data effect can be removed (not without a noise degradation) by taking the magnitude of V_1, V_2 prior to the subtractor.

The actual received signal of course arrives at the receiver at RF and has data modulation in addition. One can generate a coherent carrier for the down-conversion operation as shown in Fig. 3-3. Note that if one is accurately tracking the code the punctual channel output contains a BPSK signal. A X2 multiplier followed by a phase-locked loop can then recover the pure carrier.

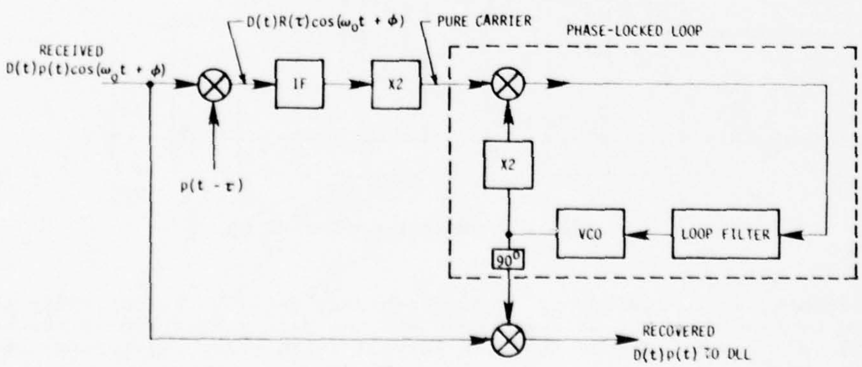


FIGURE 3-3 CARRIER RECOVERY OF THE RECEIVED SIGNAL

After carrier recovery the recovered baseband code P(t) can then be fed to the coherent delay-lock loop for code tracking. Thus if one can once get the reference code delay to match closely the received signal code delay one can recover both the carrier and the code.

An alternative to this operation is to use a noncoherent delay-lock loop as shown in Fig. 3-4. In this instance the coherent correlators of Fig. 3-2 are replaced by non-coherent bandpass correlators followed by envelope detectors. The outputs of each of the multipliers in this instance contain a narrowband bandpass signal at the IF frequency of the form

$$D(t)V_1 \qquad D(t)V_1(\tau)\cos(\omega_1 t+\emptyset) \qquad (3\text{-}1)$$

The bandpass filter (BPF) is the bandpass equivalent of the integrator (finite memory) and has sufficient bandwidth to pass only the combination of

- Residual doppler frequency uncertainty or frequency drift
- Data modulation

The IF bandwidth should be made narrow enough to reject as much of the noise as possible. The output of the BPF is then envelope detected to remove the data modulation. Subtraction of the two components then gives an error signal as in Fig. 3-2 or something similar.

It is easily seen that the noise performance of these loops is improved by decreasing the loop bandwidth to as small a value as possible. However, beyond a certain limit this improvement causes a serious degradation in the dynamic tracking performance of the loop.

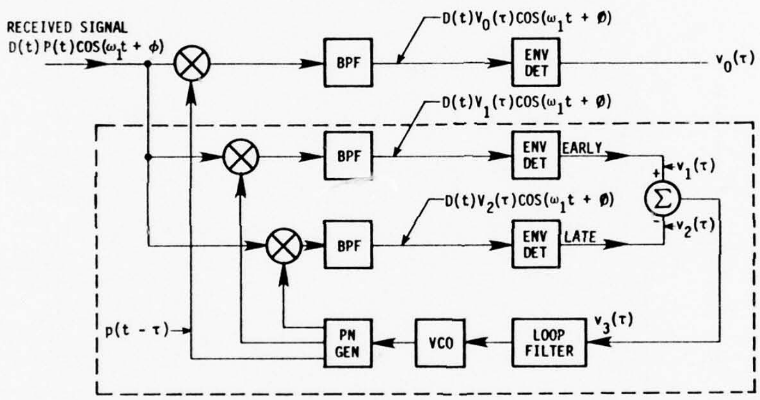


FIGURE 3-4 NONCOHERENT DELAY LOCK LOOP (DLL)

τ Dither Delay Lock Loop

The τ dither loop shown in Fig. 3-5 is a useful variation of the delay lock loop. In this variation the early and late channels are processed in time sequence rather than in parallel. The dither or time multiplexing rate must be large compared to the overall loop bandwidth but not so large so as to widen substantially the bandpass filter.

Although the τ dither loop in general does not perform quite as well as the delay-lock loop it has an implementation advantage in that it requires only a single channel correlator and still often has good performance.

3.3 Receiver Tracking Performance

The tracking error of the receiver operating on the GPS code has two major components, transient error caused by imperfectly tracking the user dynamics, and noise error caused by thermal noise.

Transient Errors

In order to model the dynamics, we assume the receiver is on an airborne platform where the radial range has a control stick jerk from 0g to a 5g steady state acceleration. The acceleration transient is shown in Fig. 3-6.

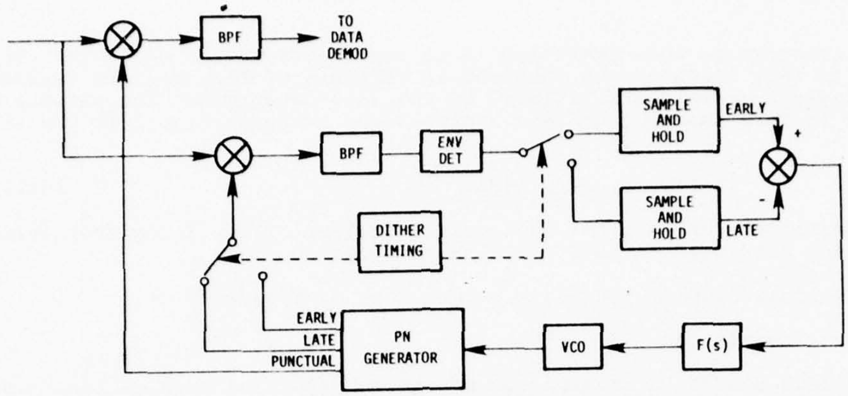


FIGURE 3-5 τ -DITHER CODE TRACKING

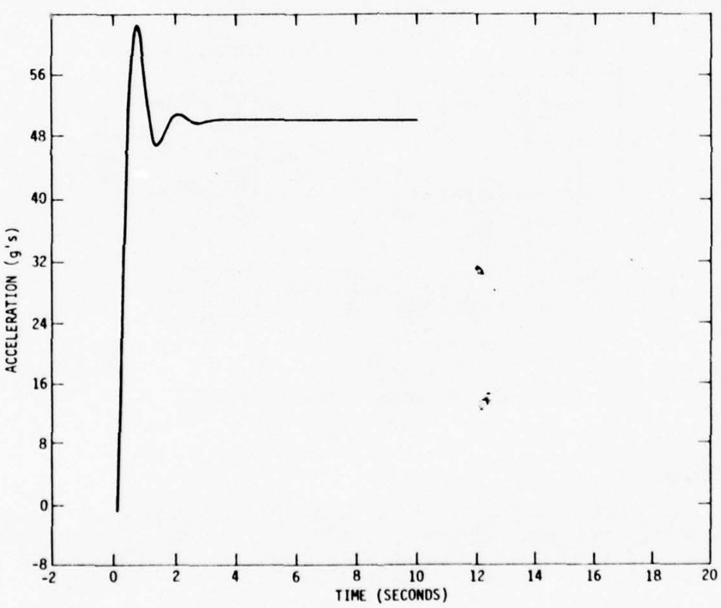


FIGURE 3-6 DRIVING FUNCTION--STEP OF STICK DEFLECTION - 5g STEADY STATE ACCELERATION WITH TRANSFER FUNCTION CORRESPONDING TO 0.8 MACH NUMBER AT SEA LEVEL

The tracking error in the delay-lock loop for 3g and 5g transients is shown in Fig. 3-7. The closed-loop bandwidth is B_L = 9 Hz and the code clock rate is 10 MHz. A second-order delay-lock loop is assumed. Figure 3-8 shows the steady state and peak transient tracking errors plotted as a function of the loop bandwidth B_L. Note that a B_L = 3 Hz yields a steady state delay error of approximately 3 and 5 nsec for 3g and 5g steady state accelerations respectively. Thus for many purposes a B_L = 3 Hz is a reasonable value.

Thermal Noise Errors

With the 3 Hz closed-loop bandwidth the tracking error caused by received thermal noise is shown in Fig. 3-9 for various values of IF bandwidths B_{IF} in the noncoherent delay-lock loop.

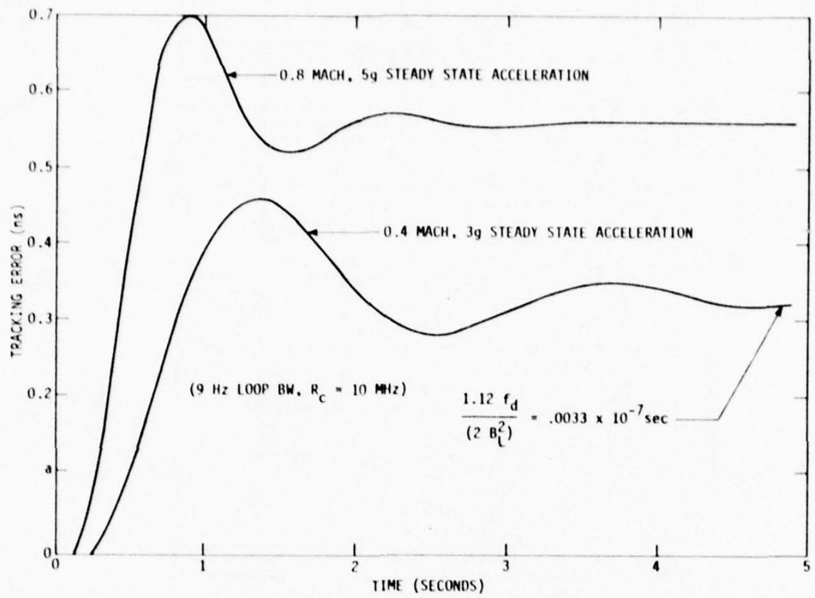


FIGURE 3-7 DLL TRACKING ERROR FOR STEP OF STICK DEFLECTION OF FIGURE 3-6

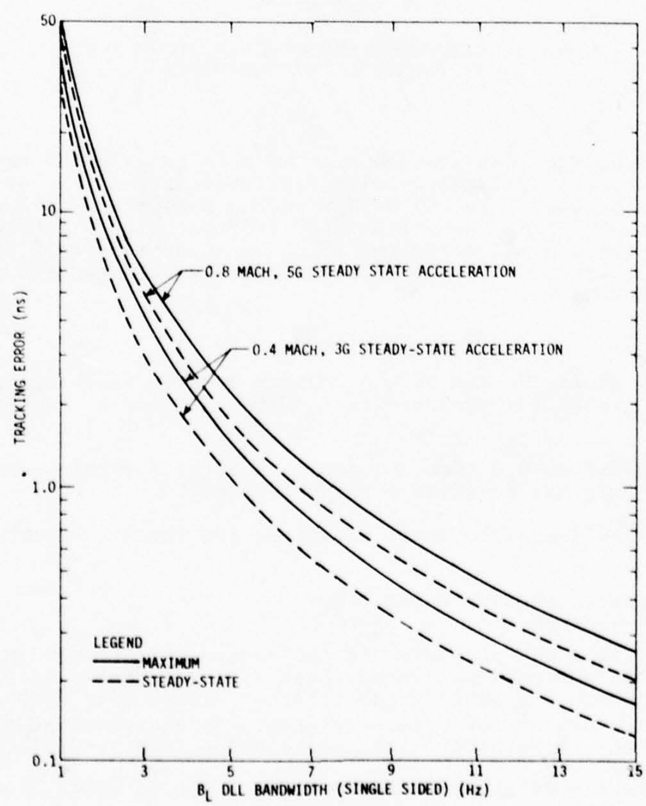


FIGURE 3-8 MAXIMUM AND STEADY-STATE DYNAMIC TRANSIENT ERROR
(2nd ORDER LOOP, R_c = 10^7 Hz)

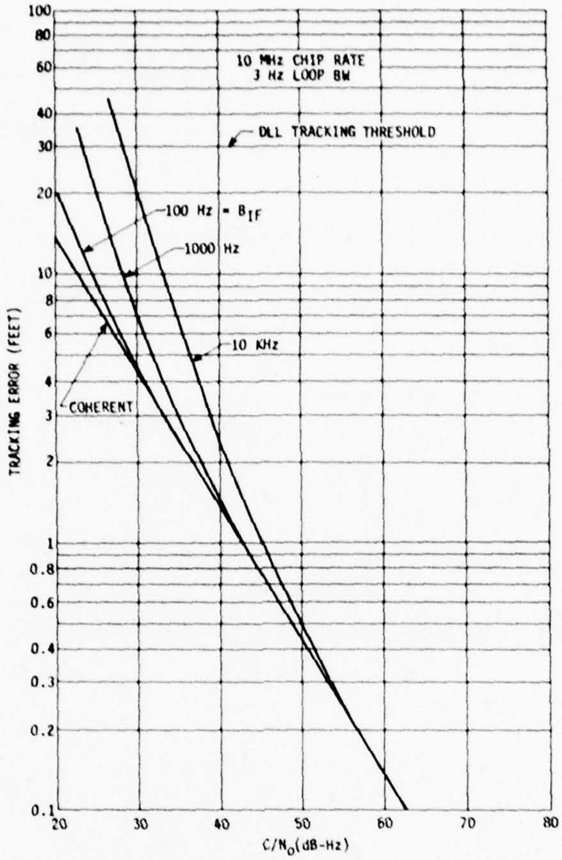


FIGURE 3-9 NOISE TRACKING ERROR IN THE DELAY LOCK LOOP VS RECEIVED CARRIER-TO-NOISE DENSITY RATIO

As already discussed the IF bandwidth must be able to pass the data modulation undistorted and tolerate any residual doppler frequency offset. Since the data modulation is 50 Hz, the PSK data bandwidth is 100 Hz and adding a doppler residual might give an bandwidth on the order of 1 kHz. Note from Fig. 3-9 that with a carrier-to-noise density ratio C/N_o = 30 dB-Hz the thermal noise rms tracking error is only 7 ft. This error compares with a 4.5 ft rms error if the IF bandwidth had been reduced to 10 Hz or a coherent loop had been employed.

Combined Tracking Error

Finally Fig. 3-10 shows the sum of the maximum dynamic tracking error ϵ_D plus the rms thermal noise error σ_T vs B_L for various C/N_o. Note that for B_L = 3 Hz this error is $\epsilon = \epsilon_D + \sigma_T = 10$ ft.

If one assumes a PDOP of 3.0 then the position error for this condition is 30 ft and we have satisfied our original position error objective.

Table 3-1 summarizes the performance equations for the second-order code tracking loop.

3.4 Received Carrier/Noise Density Ratio C/N_o

The received C/N_o is a key parameter in the system performance analysis. The received signal level to a 0 dBIC antenna has already been stated. The noise density of the received signal is kT_{eq} where k is Boltzmanns constant -198.6 dBm/°k-Hz and T_{eq} is the equivalent noise temperature of the receive system. To the received signal level we must add losses in the antenna to receiver link, correlation loss, etc.

An example calculation of the received C/N_o is given in Table 3-2.

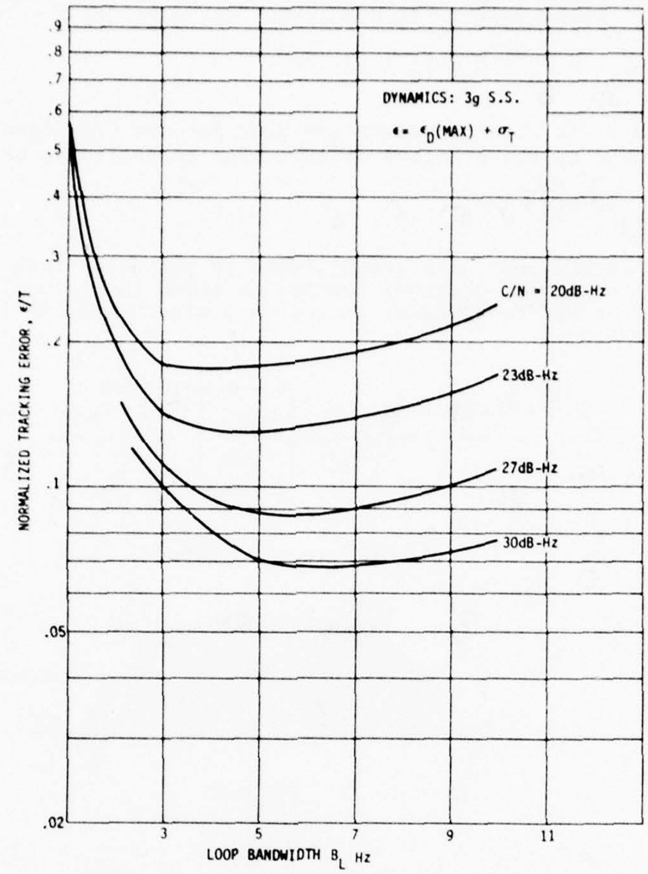


FIGURE 3-10 COHERENT DLL TRACKING ERROR DUE TO DYNAMICS AND NOISE

Table 3-1 DELAY-LOCK LOOP PERFORMANCE SUMMARY

<u>DYNAMICS</u>	$\tau_{ss} = \frac{1.12\, \tau_d\, T_c}{4\, B_L^2}$	FOR CONSTANT DOPPLER RATE, τ = DOPPLER RATE, T_c = chip period. (ACCELERATION)
<u>NOISE</u>	$\sigma_\tau^2 = \frac{N_o\, B_L}{2P_s}\left[1 + \frac{2}{SNR_I}\right]$	
<u>THRESHOLD</u>	$\frac{\tau_{ss}}{T_c} + 3\sigma_\tau = .9$	

Table 3-2 Typical received C/N_o calculation for the C/A signal

Received carrier power (C/A)	-160 dBw-Hz
Cable/Filter Losses	-1 dB
Correlation Loss	-1 dB
Net Effective C	-162 dBw
k Boltzmanns Constant	-228.6 dBw/^{o}k-Hz
T_{eq} Equivalent Noise Temp. ^{o}K	+28 dB
N_o	-200.6 dBw/^{o}k-Hz
C/N_o	38.6 dB-Hz

The system noise temperature T_{eq} is related to the antenna temperature T_a, antenna system loss, ambient temperature T_o, and receiver noise figure F by the equation

$$T_{eq} = \frac{T_a}{L} + \frac{L-1}{L} T_o + (F-1)T_o$$

This C/N_o = +38.6 dB-Hz is then the effective C/N_o for the C/A signal in this example. This result can be used in the previous calculations to determine performance and performance margin.

3.5 Data Demodulation

Once the code tracking has been accomplished by the delay-lock loop, the BPSK data at 50 bps can be recovered in the punctual channel as shown in Fig. 3-11. The received signal, either the P or C/A code signal is fed to a mixer where it is correlated with the punctual code p(t-τ) sin ωt.

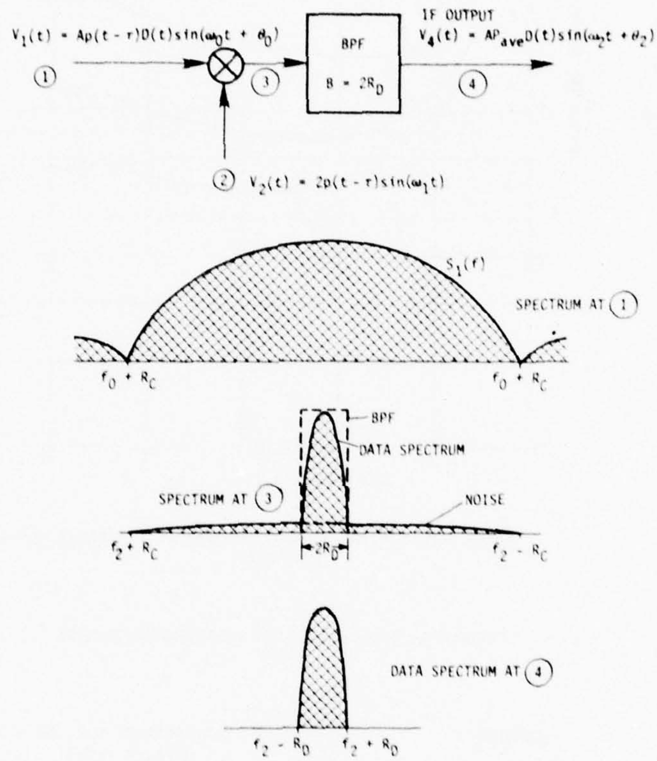


FIGURE 3-11 DATA DETECTION
(BIPHASE DATA MODULATION)

The output of this mixer/bandpass correlator is then the BPSK data signal plus additive thermal noise. This signal is then demodulated by a conventional BPSK demodulator as shown in Fig. 3-12. The filtered IF signal is first fed to either a X 2 multiplier or a Costas loop as shown. The recovered coherent carrier then is mixed with the IF signal and the baseband output is the 50 bps data stream plus noise. Data detection is then performed by bit synchronization and data detection with an integrate-and-dump filter or comparable data detector.

Figure 3-13 shows a simplified receiver block diagram for a GPS signal C/A code. The block diagram shows both the code tracking and carrier tracking/data demodulation functions. Not shown are the code acquisition, data demultiplexing and P-code handover functions. (For a detailed discussion of performance of BPSK data detection see Spilker, 1977, Chapters 11, 12.)

3.6 Search and Acquisition of GPS Codes

The tracking performance discussion of the GPS signals has assumed that somehow the reference code tracking error has been decreased to less than +1 code chip error. Initially however the user receiver may have little knowledge of its exact position and there may be a significant uncertainty as to the relative doppler offset. With the C/A code however there are a limited number, 1023, of code chips in the period; hence even with no initial knowledge of position relative to the satellite, one need only search a maximum of 1023 code chips. The maximum residual doppler uncertainty is a combination of the satellite radial velocity and the user velocity minus any prediction of these parameters available to the user.

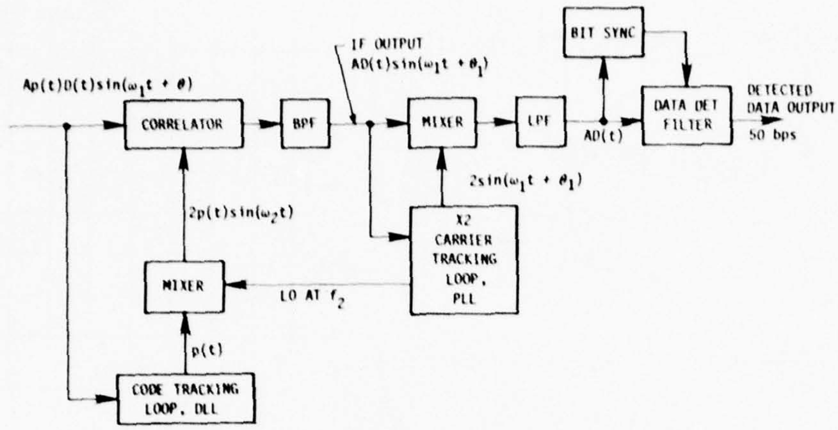


FIGURE 3-12 CODE STRIPPING AND DATA DEMODULATION

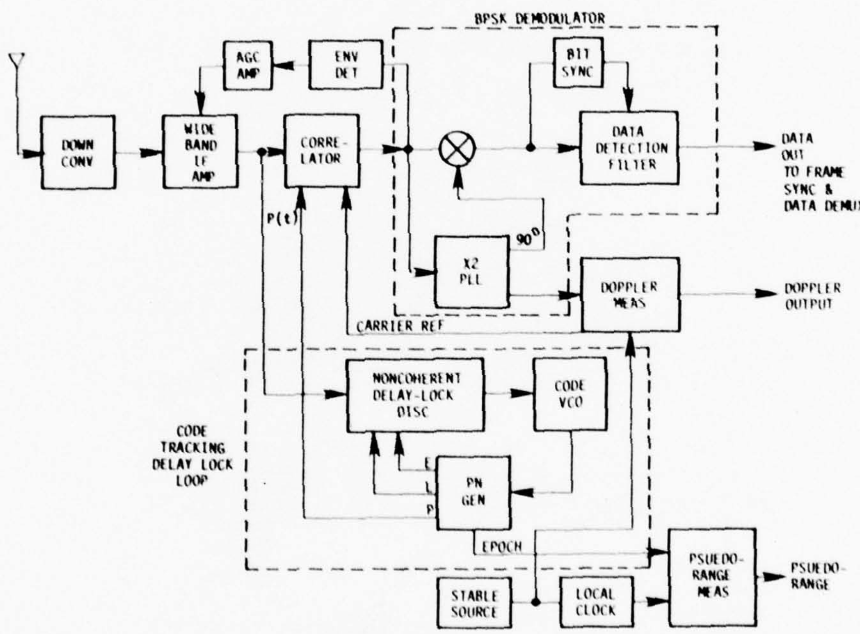


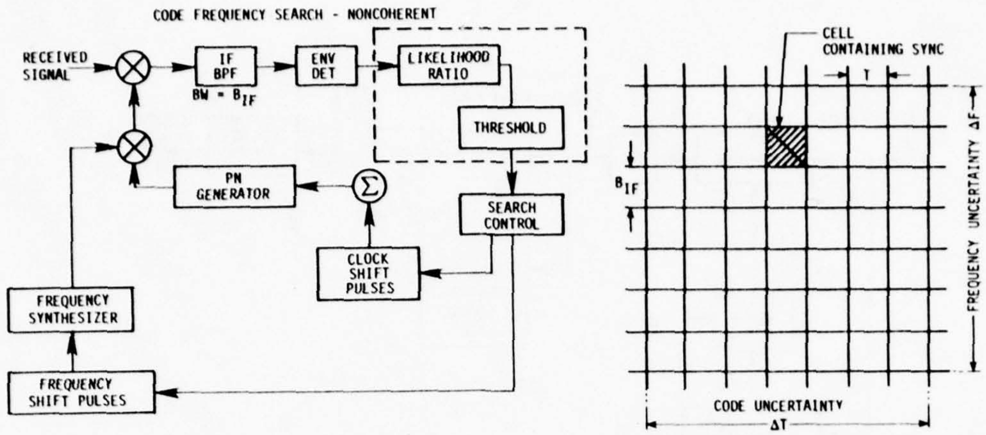
FIGURE 3-13 SIMPLIFIED BLOCK DIAGRAM OF A GPS SPREAD SPECTRUM RECEIVER

Figure 3-14 shows the time-frequency uncertainty region. Each cell has a width of 1 chip in time and a frequency width of 1 IF bandwidth. As an example the C/A chip width is approximately 1 μsec and the IF bandwidth may be on the order of 1 kHz. The overall time uncertainty ΔT is then 1023 μsec and the frequency uncertainty may be perhaps 10 kHz. Thus there are some 10^4 time-frequency cells to search. One must also account for the effect of doppler on the time cell. If doppler is present significantly, the signal time cell is itself changing with time.

Figure 3-14 also shows a typical noncoherent code-frequency search detector. The search operation essentially scans all frequency cells in the uncertainty region, measures the output power in the IF bandwidth, and integrates the power and compares with a fixed or variable threshold. When a cell is found which exceeds threshold, and it can be checked, e.g. 3 out of 4 times, then an in-lock condition is declared and search is stopped.

Figure 3-15 shows the power spectra for the noncoherent lock detector where P_s represents the power in the bandwidth compressed signal.

Code search must really be performed in half chip increments in order to avoid missing the correlation peak as shown in Fig. 3-16. If one searches in half chip increments then even at worst the cross-correlation is at least 0.75 of its maximum voltage.



SEQUENCE OF OPERATIONS

1) SENSE "OUT-OF-LOCK" CONDITION AND BEGIN CODE AND FREQUENCY SEARCH
2) SEARCH CONTINUOUSLY OR IN STEPS (FRACTION OF A CHIP) FOR SYNC
3) SENSE CODE SYNCHRONIZATION AND HALT SEARCH
4) BEGIN CARRIER RECOVERY IF COHERENT DATA DEMODULATION IS TO BE UTILIZED
5) SENSE CARRIER ACQUISITION AND STOP CARRIER SEARCH

FIGURE 3-14 ACQUISITION OF SPREAD SPECTRUM SIGNALS

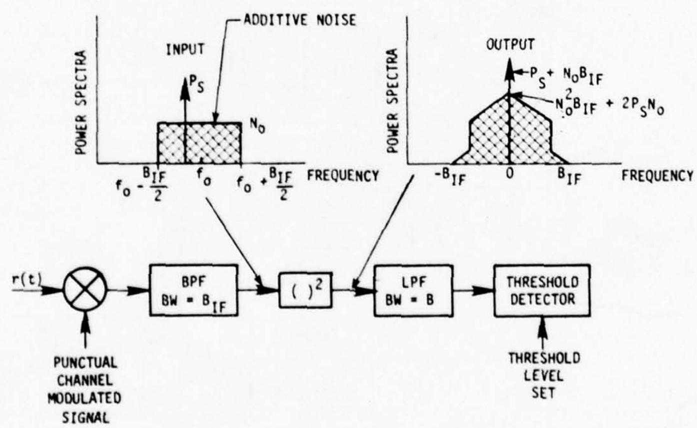


FIGURE 3-15 NONCOHERENT IN-LOCK DETECTOR - POWER SPECTRA

For a probability of detection P_D = 0.9 and a false alarm probability 0.005, the search rate, S_R, in cells/sec is approximately

$$S_R = C/22N_o \text{ with a IF bandwidth of } 22S_R. \qquad (3\text{-}2)$$

Thus if C/N_o = 33 dB or 2000 then S_R = 90 or 45 chips/sec. The doppler uncertainty tolerated is approximately $\pm$ 1 kHz or 2 kHz total. Thus it takes approximately 22 sec to search the entire C/A code period with a single 2 kHz frequency segment. Search of additional frequency uncertainty, e.g. $\pm$ 4 kHz would take 88 sec. Thus acquisition of the C/A code even with significant frequency uncertainty can be accomplished in 90 sec max. or with an expected acquisition time of 45 sec.

The total acquisition time for 4 satellites can be this same value if 4 parallel receivers are employed or 4 times this large if a single receiver is time sequenced over the four satellites.

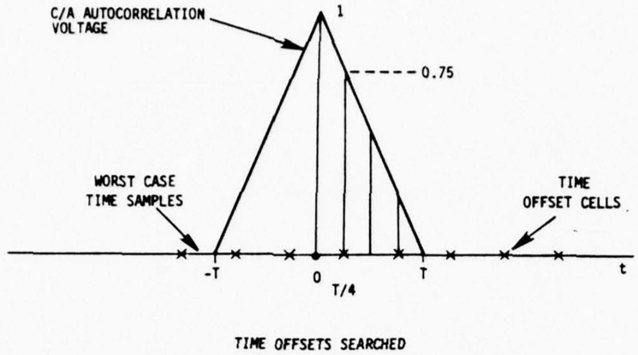


FIGURE 3-16 ACQUISITION CODE SEARCH

3.7 Multipath and Interference Effects

Multipath Reflection

Multipath signals can have a significant amplitude relative to the desired direct ray, particularly when the user is over water. Figure 3-17 shows the dependence of multipath differential delay on elevation angle θ and user altitude h. The relationship is

$$\Delta T \simeq \frac{2h}{c} \sin\theta > \frac{h}{5c} \quad \text{for } \theta > 0.1 \text{ rad.} \tag{3-3}$$

The multipath signal acts the same as noise of the same power if the delay difference ΔT is > 1.5 μsec for the C/A code or 150 nsec for the P-code and the user receiver is already tracking the desired signal.

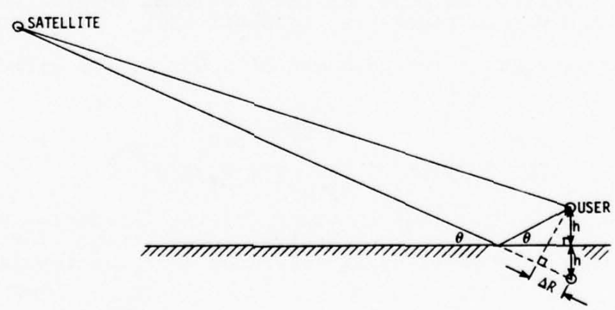


FIGURE 3-17 MULTIPATH DELAY VS ELEVATION ANGLE AND USER ALTITUDE h

Figure 3-18 shows the P-code multipath tracking error for ratios of multipath signal-to-desired signal of 0.2, 0.6, 1.0. Region 1 is the normal operating region where the signal is initially properly tracking the desired signal. Region 2 shows the performance where the receiver is offset to track the multipath rather than the desired signal. Note that for a multipath ratio of 0.6 the worst case error is only about 8 ft in Region 1.

Note that the multipath effects for the P-code are rather unlikely since one has to have a combination of low altitude and low elevation angle. For example, in order to have $c(\Delta T)$ = 150 ft $\geq \frac{h}{5}$ or h < 750 ft even at low elevation angles. For the C/A code there are

significant multipath effects even at h < 7500 ft altitude if the elevation angle is only 5^{o}. Nevertheless even for the C/A signal we have obtained a large improvement in multipath rejection relative to many alternative signals.

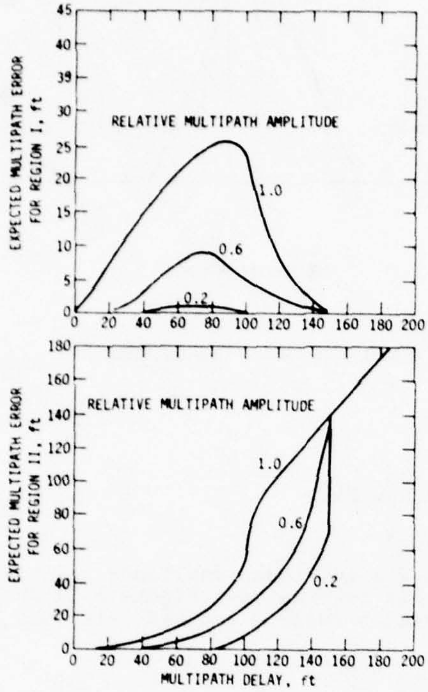


FIGURE 3-18 EXPECTED NONCOHERENT DLL TRACKING ERROR, 20 dB CARRIER TRACKING MARGIN

CW Interference Effects

Unintentional narrowband CW interferences of very low power are a likely occurrence. These signals could severely degrade any narrowband CW signal at the same frequency. However, the GPS receiver spreads the power in these CW tone interferences out over the 2 MHz or 20 MHz for the C/A and P-code receivers, respectively.

Thus if we have a narrowband interference of power P_I the effective C/N_o of the receiver is

$$(C/N_o)_{eff} = \frac{1}{(N_o/C) + (P_I/CR_c)} = \frac{C/N_o}{1 + P_I/R_c N_o} \qquad (3\text{-}4)$$

where R_c is the code clock rate. Thus in order for the interference power to be significant it must be 60 dB or 70 dB above the receiver noise density for the C/A and P-codes, respectively. Thus there is a substantial tolerance to these low level CW interferences.

REFERENCES

1. Headquarters, USAF Space and Missile Systems Organization, Systems Specification for the NAVSTAR Global Positioning System, 1 Sept. 1977.

2. R. L. Easton, L. C. Fisher, D. W. Hanson, H. W. Hellwig, L. J. Rueger, "Dissemination of Time and Frequency by Satellite," Proc. of the IEEE, October 1976, pp. 1482-1493.

3. J. J. Spilker, Jr., Digital Communications by Satellite, Prentice-Hall, Engelwood Cliffs, N.J., 1977, pp. 495-608.

4. R. Gold, "Optimal Binary Sequences for Spread Spectrum Multiplexing," IEEE Trans. on Info. Theory, Oct. 1967, pp. 619-621.

5. B. D. Elrod, "Correction for Ionospheric Propagation Delay in ASTRO-DABS - The DUAL FREQUENCY CALIBRATION METHOD," MITRE Report, MTR-6896, April 1975.

6. Hagerman, L., Private Communication, 1972.

7. A. H. Bogen, "Geometric Performance of the Global Position System," Aerospace Corporation, AD-783-210, 1974.

8. H. Chang, Private Communication, 1976.

THE GPS NAVIGATION MESSAGE

A. J. Van Dierendonck
Systems Consultant
AVAND Systems Engineering
6651 Edmonton Ave.
San Diego, California 92122 USA

S. S. Russell
Senior Engineer
General Dynamics Electronics Division
P. O. Box 81127
San Diego, California 92138 USA

E. R. Kopitske
Staff Engineer
Magnavox Research Laboratories
2829 Maricopa St.
Torrance, California 90503 USA

SUMMARY

The GPS Users continuously receive navigation information from the GPS Space Vehicles in the form of data bits modulated on the received signals. This information, which is computed and controlled by the GPS Control Segment, includes the vehicle's time, its clock correction and ephemeris parameters, almanacs and health for all GPS Space Vehicles, and text messages. From this information the Users compute the Space Vehicle's precise position and clock offset and less precise positions and clock offsets of Space Vehicles yet to be acquired.

The GPS Navigation Message design process included numerous trade studies which weighed various representations and algorithms against variables such as message size, accuracy, update frequency, User computational requirements, and legacy for the operational GPS. Other factors such as graceful degradation and future User requirements were also considered. Finally, upon selecting the appropriate design structure, the design was fine tuned to its final form and User algorithm implementation trade-offs were performed. The representation algorithms and User algorithms were jointly tested using a simulated Space Vehicle ephemeris trajectory and Space Vehicle clock. The results of these tests demonstrate that the User models represent the simulated ephemeris and clock to within 0.01 meters with precise parameters, and to within 0.1 meters with truncated parameters.

I. INTRODUCTION

The GPS Navigation Message is the information supplied to the GPS Users from a GPS Space Vehicle. It is in the form of a 50 bit per second data bit stream that is modulated on the GPS navigation signals. This signal data allows the User to navigate successfully with the GPS. It carries Space Vehicle ephemerides, system time, Space Vehicle clock behavior data, transmitter status information and C/A (Clear/Acquisition) to P (Precision) signal handover information. The data stream is common to both the P and C/A signal on both the L_1 and L_2 frequencies.

The data message is contained in a data frame that is 1500 bits long. It has five subframes, each of which contains system time and the C/A to P handover information. The first subframe contains the Space Vehicle's clock correction parameters and ionospheric propagation delay model parameters. The second and third subframes contain the Space Vehicle's ephemeris. The fourth subframe contains a message of alphanumeric characters. The fifth subframe is a cycling of the almanacs of all Space Vehicles (one per frame) containing their ephemerides, clock correction parameters and health. This almanac information is for User acquisition of yet to be acquired Space Vehicles.

The purpose of this chapter is to present the design of this GPS Navigation Message, the rationale behind the design and its impact on the Users.

II. GPS NAVIGATION MESSAGE REQUIREMENTS

The User Navigation Solution

A GPS User three-dimensional navigation fix requires pseudo-range measurements from four GPS Space Vehicles, with time being the fourth solution variable. The concept of these pseudo-range measurements is simplified and illustrated in Figure 1. These measurements are defined as the transit times of the Space Vehicle's generated signals as observed by the User and scaled by the speed of light (c). Using GPS time as a reference, the true transit times are those between the GPS transmit times and the GPS receive times. They represent the true slant ranges except for propagation delays,

described by

$$R_i = c\,(t_R - t_{T_i}) - c\,\Delta t_{A_i} \;;\; i = 1, \ldots, 4 \tag{1}$$

where t_R is GPS receive time, assumed to be simultaneous from all Space Vehicles i, t_{T_i} are the GPS times of transmission and Δt_{A_i} are the propagation delays.

The corresponding pseudo-ranges are

$$\tilde{R}_i = R_i + c\,\Delta t_{A_i} + c\,(\Delta t_u - \Delta t_{s_i}) = c\,(UT - t_{Ts_i}) \;;\; i = 1, \ldots, 4 \tag{2}$$

where Δt_u is the User's clock offset from GPS time, the Δt_{s_i} are the Space Vehicle i clock offsets, UT is User receive time defined as

$$UT = t_R + \Delta t_u \tag{3}$$

and t_{Ts_i} are the Space Vehicle i transmit times defined as

$$t_{Ts_i} = t_{T_i} + \Delta t_{s_i} \;;\; i = 1, \ldots, 4 \tag{4}$$

The User must solve for four unknowns. These are his position coordinates X, Y, and Z (earth-fixed earth-centered) and his clock offset Δt_u. Expanding Eq. (2) expressing the pseudo-ranges in terms of these unknowns yields

$$\tilde{R}_i = \sqrt{(X_{s_i} - X)^2 + (Y_{s_i} - Y)^2 + (Z_{s_i} - Z)^2} + c\,\Delta t_{A_i} + c\,(\Delta t_u - \Delta t_{s_i}) \;;\; i = 1, \ldots, 4 \tag{5}$$

There are twenty other unknowns in these four equations that must be defined before X, Y, Z and Δt_u can be found.

The Δt_{A_i} are estimated by the User by measuring the pseudo-ranges at two frequencies (ionosphere delay corrections) and estimating troposphere delays based on geometry and altitude. The X_{s_i}, Y_{s_i}, Z_{s_i}, and Δt_{s_i} must be computed by the User from information provided to him via the GPS Navigation Message. This is one purpose of the GPS Navigation Message.

<u>Space Vehicle Signal Acquisition</u>

If a User has an a priori knowledge of the pseudo-range to a Space Vehicle to within about 10,000 to 20,000 meters, he can acquire the Space Vehicle's "P" signal <u>directly</u> in a specified amount of time. <u>Direct</u> P code acquisition is possible for the User who has a <u>stable time</u> reference and has an <u>approximate</u> knowledge of his position. It is useful if he doesn't have time to <u>normally</u> acquire the Space Vehicle's Clear Acquisition "C/A" signal so as to handover to P code <u>tracking. It is</u> also useful for reacquisition or when the User is navigating and wishes to change Space Vehicles. In fact, <u>direct</u> acquisition is the normal mode in these cases.

Equation (5) will yield this approximate pseudo-range provided that the User has a knowledge of his position and time as well as the knowledge of the Space Vehicle's position and time. The atmospheric delays could be neglected. The Space Vehicle's position and time must be computed <u>before</u> acquisition. Therefore, the information for this computation must be collected from other Space Vehicles or from the same vehicle at a previous time. This information is provided to the User via the GPS Navigation Message. Each Space Vehicle provides an almanac for <u>all</u> Space Vehicles. This is another purpose of the GPS Navigation Message.

<u>Specified Requirements</u>

The GPS System Level Specification(1) specifies the signal data of the GPS Navigation Signal Structure. This specification was the basis for the GPS Navigation Message design. However, it has evolved through Specification Change Notices as the message design has evolved.

Originally, the specification was written in fairly general terms with respect to some of the contents of the data, while quite specific on other contents. It is not surprising, however, that only the more general specifications still apply. The first sentence is still intact except for an addition of the word "shall". It states: "The signal data shall allow the User to navigate successfully with GPS." Of course, this statement is <u>the</u> top level requirement. All the other requirements of the original specification were only <u>guidelines</u> for the design along with some constraints imposed by completed signal structure or Space Vehicle design. These constraints were:

1) The data rate shall be 50 bps.

2) Each data frame shall have lengths of 600, 900, 1200, 1500, or 1800 bits. (It is now fixed at 1500 bits.)

3) Each data frame shall contain handover words (HOW), and telemetry words (TLM). The HOW and TLM shall be generated on-board the Space Vehicle. Parity bits shall be generated by the Space Vehicle. Each HOW shall be spaced 6 seconds apart uniformly throughout the frame and shall contain system time. The TLM shall be used to indicate the status of the data uploading operation while it is in progress. A synchronization pattern and a within-frame identifier shall be in the HOW and/or TLM words. (The synchronization pattern was specified to be in the HOW, but ended up in the TLM.)

4) The total number of bits required in Space Vehicle memory shall be approximately 100 K bits.

There were many other "thought to be" constraints that were changed during the design process. As a result, the current system specification reflects "derived" requirements rather than specified requirements. This set of "derived" requirements is really the Navigation Message design presented here.

III. THE TLM AND HOW WORDS

Originally, the content of the TLM and HOW, their order of appearance (HOW first, followed by TLM) and their size were specified. The HOW was specified at 56 bits and the TLM at 24 bits for a total of 80 bits not including parity. They have been compacted to two 24 bit words plus parity, thus allowing more space in the message for navigation data.

The data frame is now made up of five subframes of 300 bits each. Each subframe consists of 10 words of 30 bits each. The first two words are the TLM and HOW, which are generated by the Space Vehicle. The remaining eight words of each subframe are User navigation data generated by the Control Segment.

The TLM contains an eight bit preamble (the synchronization pattern), 14 bits of TLM message, 2 non-information bearing bits and 6 bits of parity. The TLM message contains, at the appropriate time, primary upload status messages, diagnostic messages, and other messages, including roll momentum dump Z-count (Space Vehicle time of roll momentum dump).

The HOW contains a 17 bit Z-count (Space Vehicle time at the leading edge of the next subframe), a one-bit synchronization flag, a three bit subframe identification, two non-information bearing bits, and six bits of parity. The synchronization flag is an indicator to the User that the data frame may not be aligned with the X1 Code Epoch. The probability of this occurring is quite remote.

The Space Vehicle generates the parity on these two words. The Phase I Space Vehicles will not generate the parity on the other eight words; later Space Vehicles may. The parity is a (32, 26) Hamming Code, thus the parity overlaps the words. To account for this, the Space Vehicle "wastes" two bits each in the TLM and HOW so they are not linked with the other words. In later Space Vehicles, these 4 bits can contain information.

The parity algorithm will not be discussed in this chapter. It is described in detail in Reference 2.

IV. DATA BLOCK 1 - SPACE VEHICLE CLOCK CORRECTIONS

Data Block 1 appears in the first subframe, repeating itself every 30 seconds. It is generated by the Control Segment and contains frequency standard corrections, an associated Age of Data (AODC) word and ionospheric propagation delay model coefficients. The ionospheric delay model is for single frequency Users and will not be discussed in this chapter, as it is presently an experimental model. A discussion of the Space Vehicle clock corrections follows.

Space Vehicle Clock Drift Characteristics

The purpose of the clock correction parameters is to provide the Users with a description of the Space Vehicle's time offset from GPS time. This offset is not constant because the Space Vehicle's frequency standards have definite drift characteristics. These drift characteristics determine (in part) how the time offset should be represented and presented to the Users. Reference 3 presents a detailed description of GPS time and its relationship to Space Vehicle time.

During Phase I of GPS the frequency standards will be of the Rubidium type and of the Cesium beam type. The Rubidium standards exhibit deterministic drift characteristics on the order of a first derivative of frequency (drift rate) or a second derivative of phase (phase = time offset). At times in transient situations (caused by temperature variation, frequency adjustment or turn-on) they exhibit higher order derivatives. However, these higher order effects can be neglected over short intervals of time. Random drift characteristics are not predictable. They only corrupt one's ability to estimate and predict the

deterministic drift and cause some of the higher order derivatives to drift. Cesium beam standards normally don't exhibit higher order deterministic drifts in frequency over the time periods of interest. They can be characterized as a frequency offset plus random variations in frequency. Any other characteristics are only observable over very long periods of time (a day or two).

General Relativistic Effects on Space Vehicle Clocks[4, 5]

It is necessary to correct Space Vehicle time for drift due to general and special relativity. These drifts arise from the fact that the Space Vehicle clocks are located at different gravitational potentials than the GPS Users and are traveling at much higher velocities. The relativistic effects cause apparent shifts in the frequencies of the clocks.

Considering the orbits of the Phase I GPS Space Vehicles, there is the possibility of two relativity drift effects - a secular drift and a periodic drift. A large secular drift can be, and will be removed by offsetting the frequency standard frequency prior to launch. Smaller secular drift is due to an off nominal semi-major axis of the orbit. A periodic time variation is caused by a non-zero eccentricity.

User Algorithm Considerations

The model for representing the corrections for the Space Vehicle's time offset must be compatible with drift characteristics and relativistic effects described above. It must also be computationally efficient for the User (including storage requirements). This consideration would suggest the use of a polynominal expansion representation of a minimum degree, with coefficients that require a minimum multiple of 8 bit bytes. It should also represent the corrections over a long enough time interval to minimize changes in the data being transmitted, and overlapping intervals to allow time for the User to collect the change in data.

Space Vehicle Clock Correction Representation Model

A second order polynomial expansion is ideal to represent the drift characteristics of a Space Vehicle clock. It would certainly also absorb any secular relativistic effect. In fact, for reasonable time periods a first order representation would probably suffice to describe a Cesium beam frequency standard's drift, but not a Rubidium frequency standard's drift.

The primary problem in using a polynomial expansion is to include the periodic relativistic effect, which for the Phase I GPS orbits could have an amplitude of up to 70 nanoseconds (considering a maximum eccentricity of .03) with a period of approximately 12 hours. The expression for this periodic effect is

$$\Delta t_r(t) = (-4.443 \times 10^{-10} \text{ sec}/\sqrt{\text{meter}})\, e \sqrt{A} \sin E(t) \tag{6}$$

where e is the eccentricity, A is the semi-major axis and E(t) is the eccentric anomaly. A polynomial (Taylor's series) expansion of this equation for a Phase I GPS orbit with an eccentricity of .03 is approximately

$$\Delta t_r(t) \approx 6.869 \times 10^{-8} \sin E(t_o) + \left[1.002 \times 10^{-11} \cos E(t_o)\right] (t - t_o) \tag{7}$$
$$- \left[7.307 \times 10^{-16} \sin E(t_o)\right] (t - t_o)^2 - \left[3.552 \times 10^{-20} \cos E(t_o)\right] (t - t_o)^3 + \dots$$

where t_o is the reference time of the representation. t_o is at the midpoint of the time interval to maximize the length of the interval.

To maintain an accuracy of about a nanosecond, the second order term can't be neglected for a time interval longer than 0.65 hours. The third order term can't be neglected for periods over 1.69 hours. This is certainly a reasonable time interval even allowing for an overlap. In fact, the nominal period of applicability has been chosen to be one hour with one-half hour of additional applicability after the data has changed.

Thus, the representation model for the Space Vehicle clock corrections is a second order polynomial described by three coefficients a_0, a_1, and a_2 and a reference time t_{oc}.[2] More specifically, the User will correct the time received from the Space Vehicle with the equation (in seconds)

$$t = t_{SV} - \Delta t_{SV} \tag{8}$$

where

$$t_{SV} = a + a_1 (t - t_{oc}) + a_2 (t - t_{oc})^2 \tag{9}$$

where t is GPS time in seconds, t_{SV} is the Space Vehicle code phase time at message transmission time in seconds, t_{oc} is the Data Block 1 reference time in seconds, measured from the GPS time weekly

epoch (which is approximately Greenwich Mean Time Saturday night/Sunday morning midnight), and a_0, a_1, and a_2 are Data Block 1 parameters.

Note that Eqs. (8) and (9) as written are coupled. As it turns out, however, the sensitivity of Δt_{SV} to t in Eq. (9) is negligible. Thus, the User may approximate t by t_{SV} in Eq. (9). However, since GPS time spans only one week, the value of t must account for beginning or end of week cross-overs. That is, if the quantity $t - t_{oc}$ is greater than 302,400, subtract 604,800 from t. If the quantity $t - t_{oc}$ is less than -302,400, add 604,800 to t.

The parameters a_0, a_1, a_2 include all general relativistic effects on the Space Vehicle clock. The secular effects are not even corrected for by the Control Segment, since they appear to be frequency offsets, and thus absorbed in the a_1 term. The only restriction is that the a_1 term have a range large enough to absorb the effect. For other reasons it has been chosen to be an order of magnitude larger than even the nominal secular effect accounted for in the Space Vehicles prior to launch (a frequency correction of -4.484 parts in 10^{10}).

The polynomial approximation of the periodic relativistic effect does not provide a model with graceful degradation after its period of applicability. Graceful degradation may be desirable for a User if jamming prevents him from receiving new data, or if the Space Vehicle fails to transmit new data.

The User has the option, however, to subtract the correction from the coefficients, and replace it with the more exact correction of Eq. (6). That is, compute the Space Vehicle clock offset as

$$\Delta t_{SV} = (a_0 - a_{0r}) + (a_1 - a_{1r})(t - t_{oc}) + (a_2 - a_{2r})(t - t_{oc})^2 + \Delta t_r(t) \quad (10)$$

where a_{0r}, a_{1r}, and a_{2r} are

$$a_{0r} = \Delta t_r(t_{oc}) \approx K \sin M(t_{oc}) \quad (11)$$

$$a_{1r} = \left.\frac{d\Delta t_r(t)}{dt}\right|_{t = t_{oc}} \approx K\, n \cos M(t_{oc}) \quad (12)$$

$$a_{2r} = \frac{1}{2} \left.\frac{d^2 \Delta t_r(t)}{dt^2}\right|_{t = t_{oc}} \approx -\frac{K}{2} n^2 \sin M(t_{oc}) \quad (13)$$

where

$$K = (-4.443 \times 10^{-10} \text{ sec}/\sqrt{\text{meter}})\, e \sqrt{A} \quad (14)$$

$M(t_{oc})$ is the mean anomaly at time t_{oc} and n is the mean motion of the Space Vehicle. The required orbital parameters for the computations are part of Data Block 2, which will be discussed later. The angle approximations in Eqs. (11), (12), and (13) (M instead of E) are sufficiently accurate for these corrections.

Other Data Block 1 Parameters

In addition to the parameters described above, Data Block 1 also includes the Space Vehicle clock correction parameter age of data word (AODC), and $L_1 - L_2$ correction parameter for the single frequency Users (T_{GD}) and eight ionospheric correction parameters.

The age of data word (AODC) provides the User with a confidence level in the Space Vehicle clock correction. It represents the time difference (age) between the Data Block 1 reference time (t_{oc}) and the time of the last measurement update (t_L) used to estimate the correction parameters. That is,

$$\text{AODC} = t_{oc} - t_L \quad (15)$$

Based on a knowledge of the Space Vehicle frequency standard's stability and the Control Segment's ability to predict its drift, the User can model a "deweighting" function which deweights a clock correction computed with relatively old data.

T_{GD} is provided to the single frequency Users as a correction term to account for the Space Vehicle's group delay differential between L_1 and L_2. This group delay differential is the difference in the propagation times of the two signal paths prior to radiation from the Space Vehicle's antennae.

It is of no consequence to the two frequency Users because their ionospheric delay corrections are identical to that of the Control Segment. Thus, any differential is absorbed in the Space Vehicle's clock correction coefficient a_0. Because of this, it is of consequence to the single frequency User who does not observe this differential.

This differential is specified (in Reference 2) to have a mean value that will not exceed 15 nanoseconds with random variations about that mean that will not exceed 1.5 nanoseconds (one-sigma). The a_0 coefficient, however, absorbs an amplification of the mean value through the ionospheric delay correction which multiplies the differential by a factor of 1.546 (approximate). Therefore, the value of T_{GD} could be as large as 23.2 nanoseconds. The random variations will not be included in T_{GD}. They would only be part of the error budget for the a_0 coefficient (2.32 nanoseconds one-sigma).

T_{GD} is estimated by the Control Segment from long-term observations of the ionospheric delay corrections. At certain latitudes and at certain times of the day and year, the ionospheric delays are near zero. A consistent residual of these times is to be defined to be due to the group delay differential and assigned to T_{GD}.

The ionosphere correction parameters are for an experimental model for single frequency Users and will not be discussed in this chapter. They are described in detail in Reference 2.

Data Block 1 Format

Data Block 1 occupies the third through tenth thirty bit words (including parity) of the first subframe. The third and fourth words are spares. The number of bits, the scale factor of the least significant bit (LSB), which is the last bit received, the range and the units of the parameters are as specified in Table 1.

The last word in the subframe, the one containing a_0, has two non-information bearing bits, for the same reason as given for those in the TLM and HOW words. In this way, a_0 is not linked with the TLM word of the next subframe through the parity algorithm. Eventually, a_0 may have two additional bits of accuracy.

The scale factors of the LSBs were determined from the maximum sensitivities with respect to the parameters

$$\frac{\partial \Delta t_{SV}}{\partial a_0} = 1 \tag{16}$$

$$\frac{\partial \Delta t_{SV}}{\partial a_1} = t - t_{oc} \leq 2700 \text{ seconds} \tag{17}$$

$$\frac{\partial \Delta t_{SV}}{\partial a_2} = (t - t_{oc})^2 \leq 7.29 \times 10^6 \text{ seconds}^2 \tag{18}$$

The RSS of the maximum errors per term due to truncation is 0.59 nanoseconds. This is well below the expected prediction error for Δt_{SV}.

The LSB of t_{oc} has a somewhat different meaning as it is defined to be exactly GPS time. Its accuracy need only be about 0.267 seconds to provide a Δt_{SV} accurate to one nanosecond. t_{oc} is always selected to be a reference time of a multiple of 16 seconds from GPS weekly epoch prior to the definition of the coefficients. The range of t_{oc} is one week less 16 seconds since one week is the range of GPS time. The range of the coefficients are worst case expected values, or in the case of a_0, the largest value that will be allowed. Space Vehicle code phase adjustments will be performed via uploaded command from the Control Segment to maintain all Space Vehicles' times to within 976.6 microseconds of GPS time. This insures that all Space Vehicles will be transmitting the same subframe, the preamble and the same Z Count at the same time to within one twentieth of a data bit. These adjustments should be required no more than three times a year for a normal frequency standard.

The range of a_1 represents the worst case frequency offset expected, even if all redundancies have been exhausted to the point of using the crystal oscillators. Frequency adjustments via the S-band telemetry are possible, however, they will not be normally utilized.

The range of a_2 is primarily 2^7 (= 128) times the LSB. It is only 4.87 times larger than the periodic relativity effect described in Eq. (7). Drift rates of Rubidium standards should never be that large. The backup crystal oscillators may have drift rates that large.

Table 1. Data Block 1 Parameters

Parameter	No. of Bits	Scale Factor (LSB)	Range*	Units
Spare	24	--	--	--
Spare	24	--	--	--
α_0	8	2^{-31}	$\pm 2^{-24}$	seconds
α_1	8	2^{-31}	$\pm ^{-24}$	seconds
α_2	8	2^{-29}	$\pm 2^{-22}$	seconds
α_3	8	2^{-28}	$\pm 2^{-21}$	seconds
β_0	8	2^{8}	$\pm 2^{15}$	seconds
β_1	8	2^{9}	$\pm 2^{16}$	seconds
β_2	8	2^{10}	$\pm 2^{17}$	seconds
β_3	8	2^{12}	$\pm 2^{19}$	seconds
T_{GD}	8	$2^{-31} \simeq 4.66 \times 10^{-10}$	$\pm 2^{-24} \simeq \pm 5.96 \times 10^{-8}$	seconds
AODC	8	$2^{11} = 2048$	$2^{19} = 524288$	seconds
t_{oc}	16	$2^{4} = 16$	604,784	seconds
a_2	8	$2^{-55} \simeq 2.78 \times 10^{-17}$	$\pm 2^{-48} \simeq \pm 3.553 \times 10^{-15}$	sec/sec^2
a_1	16	$2^{-43} \simeq 1.14 \times 10^{-13}$	$\pm 2^{-28} \simeq \pm 3.725 \times 10^{-9}$	sec/sec
a_0	22	$2^{-32} \simeq 4.656 \times 10^{-10}$	$\pm 2^{-10} \simeq \pm 9.766 \times 10^{-4}$	seconds

*($\pm$) indicates that sign bit will occupy the most significant bit (MSB).

NOTE: All binary numbers will be two's complement.

V. DATA BLOCK 2 SPACE VEHICLE EPHEMERIS REPRESENTATION

Data Block 2 appears in the second and third subframes, repeating itself every 30 seconds. It is generated by the Control Segment and contains a representation of the transmitting Space Vehicle's ephemeris and associated Age of Data (AODE) words. A discussion of this ephemeris representation follows.

GPS Space Vehicle Orbit Characteristics

The characteristics of the GPS Space Vehicle orbits are only summarized here. A more detailed discussion appears in Reference 6.

The final orbits of the GPS Phase I Space Vehicles are characterized by an orbital ground trace which repeats once each sidereal day. These orbits have a near one-half sidereal day period. The nominal period, including nodal regression effects, will be 11.9661 hours which corresponds to a semi-major axis value of 26560.123 Km. The inclination angle will be 63 degrees and a nominal eccentricity of less than 0.005 (with an upper limit of 0.015).

Even though the orbits are nominally circular, they must be described as elliptical orbits. The radial variation in the orbit (with respect to the earth) varies from

$$r_{min} = A(1 - e) < 26427.322 \text{ Km} \quad (19)$$

to

$$r_{max} = A(1 + e) > 26692.924 \text{ Km} \quad (20)$$

for eccentricities greater than .005. This is a variation of 265.602 Km.

The eccentricity causes the greatest perturbation from a circular orbit. However, there are other perturbations from even the elliptical orbit that are significant. These are due to perturbing accelerations that will act on the GPS Space Vehicles. The approximate values of these accelerations are given in Table 2. Over a short period of time, these forces could be considered as linear perturbations to estimate their effects on the ephemeris. These effects are also listed in Table 2 for a one hour time period. Because of the nature of these effects and their magnitude, it is desirable to represent the GPS ephemerides over periods of time no longer than a few hours.

The second zonal harmonic provides the major perturbing force to the two body elliptical orbit with a dominant period of half the orbit period (that is, 5.98305 hours). There also exists some secular drift in the ephemeris.

The next most predominant forces are the gravities of the Sun and Moon. These forces vary as a function of the position of the Space Vehicle in its orbit. They are nearly constant over short intervals of time. However, when the Space Vehicle is at a position of its closest approach to the moon, there are some pronounced variations in its ephemeris.

All other forces can essentially be considered constant over short intervals of time, or they have a negligible effect. Any residuals would be reasonable representation errors.

In summary, the GPS Space Vehicle ephemerides are characterized by elliptical orbits with both periodic and secular perturbations. The predominant periodic perturbations have a 5.98305 hour period. Other perturbations are smaller and might be represented over short intervals of time as simple functions of time (i.e., constant or linear).

Table 2. Summary of the Approximate Perturbing Forces For GPS Space Vehicles

Source	Maximum perturbing acceleration m/sec^2	Maximum excursion growth in one hour meters
Earth-mass attraction	5.65×10^{-1}	---
Second zonal harmonic	5.3×10^{-5}	300
Lunar gravity	5.5×10^{-6}	40
Solar gravity	3×10^{-6}	20
Fourth zonal harmonic	10^{-7}	0.6
Solar radiation pressure	10^{-7}	0.6
Gravity anomalies	10^{-8}	0.06
All other forces	10^{-8}	0.06

Ephemeris Representation Trade Studies and Results

The final GPS ephemeris representation evolved through numerous trade studies. There were many criteria to be considered. Some of these are (not necessarily in order of importance):

1) Users Time-To-First-Fix (TTFF) - This basically is a function of how many subframes are required to transmit and collect the ephemeris representation. Does the clock and ephemeris representations data fit in two, three, or four subframes? There was a time when it was thought that two subframes would suffice.

2) User Computational Time - Complexity of algorithm to compute earth-centered earth-fixed Space Vehicle position; the number of sines, cosines, square roots, multiples, divides, etc. This is only critical if User computational through-put is a problem.

3) User Storage Requirements - A function of how many subframes of data must be stored and the computer code required for the User algorithm. There is also a requirement for an almanac algorithm to compute the positions of all Space Vehicles. Thus, commonality of algorithms impacts the storage requirements.

4) Refresh Rate - How long is one set of parameters valid before they must be replaced? This impacts the memory required in the Space Vehicle as well as how often the User must collect a new set of parameters.

5) Overlap in Time of Applicability - This directly impacts the refresh rate.

6) Accuracy - How well is the ephemeris and the rate of change of ephemeris represented? Accuracy is usually directly proportional to the number of parameters required and how many bytes are needed to store them. Almanac accuracy is also important.

7) Orbital Tolerance - How well does the representation represent off-nominal orbits? What is the range of the parameters? Orbital tolerances are hard to pin down. That is, how off-nominal can an orbit get before the Space Vehicle is no longer useful? Large tolerances require more bytes of information.

8) Degradation - Does the representation degrade gracefully after its period of applicability, or does it degrade abruptly? Degradation of representation was discussed earlier for the clock correction parameters. It would be desirable to navigate in a degraded mode with "old" parameters if new ones are not obtainable.

9) Clock Relativity Compensation - Does the representation aid graceful degradation of Space Vehicle clock corrections?

10) Time for User to Receive Almanac - Related to criteria 1 and 3. How much data is required to represent a common algorithm almanac? In order to keep the almanac data to a minimum, the almanac representation should represent multiple orbits. To do that plus have a common representation with the precise ephemeris constrains that representation considerably.

11) Clarity of the Representation - Is it clear to facilitate debugging and interpretation and can the algorithm be extended to future considerations, to other computers, etc. ? For the representation to have some physical meaning enhances the understanding of its design.

Considering the above criteria, three representations or variations thereof were investigated. Candidate representations were:

1) Polynomials in time
2) Harmonic expansions
3) Keplerian parameters plus perturbations (polynomials, harmonics)

Polynomials were considered because the User algorithm would be simple, minimizing the processing time. Algorithm storage requirements are also minimal, however, polynomials could never by themselves represent total orbits or multiple orbits. Thus, a separate algorithm would be required for the almanac computations. Thus, other than the minimal processing time, polynomial representations had no clear advantage over other candidates, and had some definite disadvantages.

Harmonic expansions around a circular orbit were also considered but not investigated in depth. This was because they didn't appear to have any clear advantages over a Keplerian type representation with the main disadvantages of not having a clear physical meaning, making it difficult to determine the range of the coefficients for data word sizing. The coefficients would be sensitive to position in an elliptical orbit and to unknown orbital tolerances.

Keplerian parameter representation provided advantages in all criteria except for User computation time and possibly, User storage requirements. The storage requirement disadvantage tends to be offset because this representation handles the almanac very well. A definite advantage was the fact that the Keplerian representation has a physical meaning - at least a familiar one. This made it relatively easy to size the data words, even for off nominal orbits. However, perturbations about the Keplerian orbit require additional parameters. Representation candidates for the perturbations were polynomials (in time) and harmonic coefficients. The harmonic coefficients have a more familiar physical meaning, therefore they provided a better representation when combined with two secular drift terms. They also provided a more graceful degradation using "old" parameters.

It should be clear by now that the choice of representation was that of Keplerian type parameters plus secular drift terms and harmonic coefficients. Table 3 summarizes the candidates versus the selection criteria described above. A detailed description of the exact choice of parameters follows.

User Algorithm Considerations[7]

Because the Keplerian representation could have had a severe impact on the User's computation time, extensive algorithm manipulations were considered to optimize the User algorithm and to determine the impact. The primary emphasis was in the solution of Kepler's equation

$$E(t) = M(t) + e \sin E(t) \tag{21}$$

Table 3. Ephemeris Representation Candidates Versus Selection Criteria

	Polynomials	Harmonic* Expansion	Keplerian with Polynomials	Keplerian with Harmonics
No. of Subframes	3+	2+	2+	2-
User Computation Time	short, simple	sines, cosines	sines, cosines	sines, cosines
User Storage Requirements	3+Subframes Small algorithm Need almanac	2+Subframes Med. algorithm Same as almanac	2+Subframes Large algorithm Same as almanac	2 Subframes Large algorithm Same as almanac
Refresh Rate	Once per hour	Once per hour	Once per hour	Once per hour or longer
Refresh Overlap	½ hour	½ hour	½ hour	½ hour or longer
Accuracy	< 1 ft.	Probably OK	< 1 ft.	< 1 ft.
Orbital Tolerance Effects	Not clear	Not clear	Handles any orbit	Handles any orbit
Degradation	Abrupt	Unknown	Marginal	Graceful
Clock Relativity Compensation	Not compatible	Not compatible	Compatible	Compatible
Almanac Subframe	Not compatible	1+	1-	1-
Clarity	Not clear	Not clear	Orbit-clear Perturbations-not clear	Clear

*Estimated characteristics

and subsequent solutions for the true anomaly

$$\sin v(t) = \frac{\sqrt{1 - e^2}\,\sin E(t)}{1 - e\cos E(t)} \tag{22}$$

$$\cos v(t) = \frac{\cos E(t) - e}{1 - e\cos E(t)} \tag{23}$$

to a defined accuracy. Other manipulations and approximations were investigated, consisting mostly of the selection of the best order of computation to minimize operations, using appropriate trigonometric identities and using small angle approximations for the sines and cosines of the harmonic perturbations.

In these equations, e is given and $M(t)$ is computed easily from given data. In general, Eq. (21) is nonlinear in $E(t)$. It is impractical to solve for $E(t)$ in any way except approximately because, for $e \le 0.663$, the exact solution is

$$E(t) = M(t) + 2\sum_{k=1}^{\infty} \frac{1}{k!} J_k(ke)\sin[kM(t)] \tag{24}$$

where the J_k are Bessel functions of the first kind of order k. Therefore, a number of methods of solution were considered in evaluating the Keplerian representation and in choosing the appropriate algorithm. Four factors were considered (in order of importance):

1. Accuracy
2. Legacy and Clarity
3. Duty Cycle
4. Memory

The guidelines followed relative to these factors were:

Accuracy. The algorithm was to introduce an error of less than 1 centimeter for eccentricities of less than 0.03.

Legacy and Clarity. The algorithm was to be suitable for implementation on digital computers with a broad range of architectures, word lengths, instruction sets and internal speeds. The algorithm

structure was to be clear to facilitate debugging and interpretation.

Duty Cycle. The execution time of the algorithm was to be minimized subject to the accuracy and legacy constraints.

Memory. Memory was not to be wasted but was to be traded for duty cycle, legacy and accuracy. If it becomes necessary to save memory at the expense of other factors, clarity was to be sacrificed.

There are two basic approaches in solving the total set of Eqs. (21) through (23). One solves for E(t) explicitly and then solves for sine and cosine of v(t), while the other solves the "equation of the center" which contains an implicit solution of Kepler's equation and yields the true anomaly v(t) directly. The "equation of the center" algorithms generally yield shorter duty cycles, where the explicit solutions generally require less memory.

There are also both iterative or closed form approximations for each approach. Closed form approximations are faster but use more memory than iterative methods for the same level of accuracy. They are also usually easier to debug, maintain, size, scale and schedule.

Ten different methods were evaluated. Six of these methods were discarded by inspection. The remaining four were tested against the four factors given above. Two of these were tested with modifications and/or different coding methods to produce results with different orderings of the factors. The results of these seven tests are summarized in Table 4. The numbers are typical for an HP21M20 computer programmed in a higher order language. They do not necessarily reflect relative efficiency in other processors and/or languages.

Table 4. Memory and Duty Cycle Requirements of Several Candidate Ephemeris Algorithm Implementation

Algorithm	Relative Order of Importance	Memory Required for Four Space Vehicles (words)	Cycle Time Per Call (milliseconds)*
1 Using Modified Lagrange Solution of Equation of Center	Accuracy Clarity Speed Memory	2203	22.3
2 Using Modified Lagrange Solution of the Equation of the Center	Accuracy Speed Memory Clarity	1496	12.5
3 Using Classic Successive Substitutions to Solve Kepler's Equations	Accuracy Memory Speed Clarity	1145	40.9
4 Using Stephenson's Successive Substitution Method to Solve Kepler's Equations	Accuracy Speed Memory Clarity	1202	46.4
5 Using Classic Newton Raphson Method to Solve Kepler's Equations	Accuracy Clarity Speed Memory	1515	55.0
6 Using Modified Newton Raphson to Solve Kepler's Equation	Accuracy Clarity Memory Speed	1755	35.9
7 Using Modifed Newton Raphson to Solve Kepler's Equation	Accuracy Memory Speed Clarity	1319	20.6

*HP21M20 Computer - FORTRAN IV

Algorithms 1 and 2 are differently coded versions of a closed form approximate solution to the "equation of the center." Basically, sine v(t) and cosine v(t) are solved as series expansions of powers of e and sines and cosines of M(t) (sixth or seventh power). Algorithm 3 consists of an iterative solution of Kepler's equation of the form

$$E_k(t) = M(t) + e \sin \left[E_{k-1}(t)\right] \tag{25}$$

where

$$E_o(t) = M(t) \tag{26}$$

until a desired accuracy is achieved (about 3 iterations).

Algorithm 4 is a variation of Algorithm 3 where first and second differences are used to accelerate convergence. Because Algorithm 3 converges so fast, Algorithm 4 was actually inferior.

Algorithm 5 uses a classical Newton Raphson method as an iterative solution to Kepler's equation by searching for an E(t) that minimizes the function

$$f\left[E(t)\right] = M(t) - \left[E(t) - \sin E(t)\right] \tag{27}$$

with

$$E_o(t) = M(t) \tag{28}$$

Algorithms 6 and 7 are simply differently coded versions of an approximation of the combined first two iterations of Algorithm 5, which was shown to be sufficiently accurate.

There were two results of this trade study. First of all, it demonstrated that the Keplerian approach to ephemeris representation was a viable approach with regard to User algorithm considerations. The memory requirement for any of the algorithms was not excessive nor was the cycle time. Considering eight calls to the routine for each navigation cycle (to evaluate range and delta-range), the total cycle time would be in the order of 100 to 200 milliseconds for this more desirable algorithm. User navigation cycles would occur every 1 to 10 seconds depending on the user set or type. Although 100 to 200 millisecond period is a good portion of a one second navigation cycle, it is not an unreasonable portion.

The other result was a selection of the most desirable algorithms. These were either Algorithms 1 or 2, depending on whether or not the clarity factor would have to be waved for better speed and less memory.

The Space Vehicle Ephemeris Representation Model

The Space Vehicle ephemeris representation model is characterized by a set of parameters that is an extension to the Keplerian orbital parameters describing the Space Vehicle ephemeris during the interval of time (at least an hour) for which the parameters are transmitted. They also describe the ephemeris for an additional interval of time (at least one-half hour) to allow time for the User to receive the parameters for the new interval of time. The definitions of the parameters are given in Table 5. The age of data word (AODE) provides the User with a confidence level in the ephemeris representation parameters. AODE represents the time difference (age) between the reference time (t_{oe}) and the time of the last measurement update (t_L) used to estimate the representation parameters. That is,

$$\text{AODE} = t_{oe} - t_L \tag{29}$$

The AODE word also provides a linkage between Subframes 2 and 3 because it appears in both subframes. This is to ensure the User that the subframes he collected do not apply to different intervals of time, as he may collect those subframes at different times.

The User computes the earth fixed coordinates of the position of the Space Vehicle's antenna phase center with a variation of the equations shown in Table 6. The parameter values are obtained via a nonlinear iterative least squares curve fit of the predicted Space Vehicle's antenna phase center ephemeris (time-position quadruples; t, x, y, z) over the interval of time $t_{oe} - T$, $t_{oe} + T$. T is 5 minutes longer than half the transmission interval plus half the overlap period into the next interval. For example, if the transmission interval is one hour, and the overlap period is 30 minutes, T is 50 minutes.

The equations given in Table 6 provide the Space Vehicle's antenna phase center in earth-centered earth-fixed Cartesian coordinates. The system is characterized as follows:

Table 5. Ephemeris Representation Parameters

M_o	Mean anomaly at reference time
Δn	Mean motion difference from computed value
e	Eccentricity
$\sqrt{A}$	Square root of the semi-major axis
Ω_o	Right ascension at reference time
i_o	Inclination angle at reference time
ω	Argument of perigee
$\dot{\Omega}$	Rate of right ascension
C_{uc}	Amplitude of the cosine harmonic correction term to the argument of latitude
C_{us}	Amplitude of the sine harmonic correction term to the argument of latitude
C_{rc}	Amplitude of the cosine harmonic correction term to the orbit radius
C_{rs}	Amplitude of the sine harmonic correction term to the orbit radius
C_{ic}	Amplitude of the cosine harmonic correction term to the angle of inclination
C_{is}	Amplitude of the sine harmonic correction term to the angle of inclination
t_{oe}	Ephemeris reference time
AODE	Age of Data (Ephemeris)

1) x is in the true equatorial plane in the direction of the Greenwich meridian
2) z is along the true earth spin axis, positive in the northern hemisphere
3) y completes the right hand system

$$y = (z) \times (x) \tag{30}$$

It must be emphasized that these representation parameters are the result of a curve fit and are only Keplerian in appearance. They only describe the ephemeris over the period of applicability and not for the total orbit, however, they do describe the true Keplerian orbit to within a few thousand meters.

There are usually only six Keplerian parameters if the reference time is the time-of-perigee. However, in this case seven are needed (M_o, t_{oe}, e, $\sqrt{A}$, Ω_o, i_o, ω), replacing the time-of-perigee with the pair (M_o, t_{oe}). The net effect of this replacement requires no increase in the number of words. This results because of the decrease in sensitivity of the time derivative parameters since t_{oe} is always within a few minutes of the evaluation time. It also improved the stability of the curve fit algorithm for near circular orbits.*

---

*Incidently, a least squares curve fit algorithm would be ill-conditioned for a perfectly circular orbit with no perturbations because M_o is redundant with ω. However, there is no such thing as a perfectly circular orbit other than instantaneously. The algorithm converged for values of e as small as 10^{-8}, although it converged to negative values of e depending on the initial selection of M_o and ω. This condition was easily corrected by adding (or subtracting) π to both M_o and ω and then converging to a positive e. Eccentricities of less than 10^{-8} could not be generated over the appropriate time intervals. If e were zero, either M_o or ω could be fixed, while letting the other vary in the fit.

Table 6. Ephemeris Representation Definitions

$\mu = 3.986008 \times 10^{14} \frac{\text{meters}^3}{\text{sec}^2}$	WGS 72 Value of the Earth's universal gravitational parameter
$\dot{\Omega}_e = 7.292115147 \times 10^{-5} \frac{\text{rad}}{\text{sec}}$	WGS 72 Value of the Earth's rotation rate
$A = (\sqrt{A})^2$	Semi-major axis
$n_o = \sqrt{\frac{\mu}{A^3}}$	Computed mean motion
$t_k = t - t_{oe}$*	Time from epoch
$n = n_o + \Delta n$	Corrected mean motion
$M_k = M_o + nt_k$	Mean anomaly
$M_k = E_k - e \sin E_k$	Kepler's equation for eccentric anomaly
$\cos v_k = (\cos E_k - e) / (1 - e \cos E_k)$ $\sin v_k = \sqrt{1 - e^2} \sin E_k / (1 - e \cos E_k)$	True anomaly
$\phi_k = v_k + \omega$	Argument of latitude
$\delta u_k = C_{us} \sin 2\phi_k + C_{uc} \cos 2\phi_k$	Argument of Latitude Correction (2nd Harmonic Perturbations)
$\delta r_k = C_{rc} \cos 2\phi_k + C_{rs} \sin 2\phi_k$	Radius Correction (2nd Harmonic Perturbations)
$\delta i_k = C_{ic} \cos 2\phi_k + C_{is} \sin 2\phi_k$	Correction to Inclination (2nd Harmonic Perturbations)
$u_k = \phi_k + \delta u_k$	Corrected argument of latitude
$r_k = A(1 - e \cos E_k) + \delta r_k$	Corrected radius
$i_k = i_o + \delta i_k$	Corrected inclination
$x'_k = r_k \cos u_k$ $y'_k = r_k \sin u_k$	Positions in orbital plane
$\Omega_k = \Omega_o + (\dot{\Omega} - \dot{\Omega}_e) t_k - \dot{\Omega}_e t_{oe}$	Corrected longitude of ascending node
$x_k = x'_k \cos \Omega_k - y'_k \cos i_k \sin \Omega_k$ $y_k = x'_k \sin \Omega_k + y'_k \cos i_k \cos \Omega_k$ $z_k = y'_k \sin i_k$	Earth fixed coordinates

* t is GPS system time at time of transmission, i.e., GPS time of reception corrected for transit time (range/speed of light). Furthermore, t_k must be the actual total time difference between the time t and the epoch time t_{oe}, and must account for beginning or end of week crossovers. That is, if t_k is greater than 302,400, subtract 604,800 from t_k. If t_k is less than -302,400 add 604,800 to t_k.

The additional parameters in Table 5 do not adequately describe an ephemeris with its many perturbations over long periods. They do, however, to the required accuracy over intervals of 1.5 to 5 or more hours. They primarily describe the second zonal harmonic effects while absorbing all other effects during those intervals of time. Table 7 lists the major contributions that justify each of the additional parameters.

These additional parameters add very little to the User's computation requirement because they either add to total angles before sines or cosines are computed; or, if they represent angle perturbations, it suffices to use small angle approximations.

The figure of merit used to measure the quality of the curve fit is User Equivalent Range Error (UERE). UERE is the projection of the curve fit error onto the User's range. The curve fit results in a UERE of less than .01 meter, one sigma. This is, of course, considerably less than the ephemeris prediction error, which is specified to be less than 12 feet in UERE (3.658 meters) one sigma, for the Phase I GPS or 5 feet in UERE (1.524 meters) one sigma for the Phase III GPS.

The UERE described above is for the results of an algorithm where the parameters are double precision floating point numbers (32 bit machine) and only represent the curve fit algorithm errors. The parameters are, of course, truncated for the GPS Navigation Message. The effect of that truncation increases the one sigma error to 0.1 meters over the time period of the fit. As stated earlier, it is desirable that the representation also degrades gracefully after its period of applicability. The parameters described in the preceeding paragraphs exhibit a reasonable degradation if used beyond their period of applicability. In fact, it is exceptional relative to other candidate representations. The relationship between the curve fit error, the truncation error and the degradation error is presented in Figure 2.

Table 7. Contributions Justifying Parameters in Addition to Keplerian Parameters

Parameter	Justification[38]
Δn	For near circular orbits, there is very little difference between Δn and dM/dt over the time interval of interest. Absorbing the effects of either in Δn, rather than in dM/dt, simplifies the User's algorithms for computing time and orbit computations. Primarily secular drift in dM/dt due to the second zonal harmonic. Also absorbs effects of sun and moon gravity and solar radiation pressure over the interval of fit.
$\dot{\Omega}$	Primarily secular drift in right ascension due to the second zonal harmonic. $\dot{\Omega}$ also includes effect of earth wobble and polar wander.
C_{uc}, C_{us}, C_{rc}, C_{rs}, C_{ic}, C_{is}	Second zonal harmonic. The theoretical solutions to Lagrange's equations for the J_2 perturbations do not always indicate the necessity of both the cosine and sine terms. However, experimentally they were quite useful as they account for nonzero phase angles introduced due to perturbations attributed to other than second zonal harmonics. The parameters also absorbed short-term effects of the moon's gravity during the closest approach of the Space Vehicle to the moon.

Data Block 2 Format

Data Block 2 occupies the third through the tenth thirty bit words (including parity) of the second and third subframes. The number of bits, the scale factor of the least significant bit (LSB), which is the last bit received, the range and the units of the parameters are as specified in Table 6. The last word of subframe 2 is a six bit spare word and the last word of subframe 3 is a 14 bit spare word. They are both part of the tenth thirty bit word of a subframe that has two non-information bearing bits. Thus, it does not link with the TLM word of the next subframe through the parity algorithm.

The scale factors of the LSBs were determined through a sensitivity analysis. This analysis computed the sensitivity of the User's range to the Space Vehicle due to truncation of the ephemeris parameters. Defining the truncation as a vector

$$e_T = x^* - x_T \tag{31}$$

where x^* and x_T are vectors of ephemeris parameters without and with truncation respectively. At a given time, the User's range R to the Space Vehicle can be expanded in a Taylor's series as

$$R = R^* + \left.\frac{\partial R}{\partial x}\right|_{x^*} e_T + O\left[e_T^2\right] \tag{32}$$

or

$$\delta R = R - R^* \approx \left.\frac{\partial R}{\partial x}\right|_{x^*} e_T \tag{33}$$

It was required that the range error due to truncation be less than 0.3 meters. That is

$$\left.\frac{\partial R}{\partial x}\right|_{x^*} e_T < 0.3 \text{ meters} \tag{34}$$

It was assumed that the contribution to range error is uniform over the fourteen components (e_{T_j}) of e_T. The jth contribution of inequality 34 is then

$$\left.\frac{\partial R}{\partial x_j}\right|_{x^*} e_{T_j} < \frac{0.3}{14} \text{ meters} \tag{35}$$

Table 8. Data Block 2 Parameters

Parameter	No. of Bits	Scale Factor (LSB)	Range*	Units
AODE	8	$2^{11} = 2048$	524,288	seconds
C_{rs}	16	$2^{-5} = .03125$	± 1024	meters
Δn	16	$2^{-43} \approx 1.14 \times 10^{-13}$	$\pm 3.73 \times 10^{-9}$	semicircles/sec
M_o	32	$2^{-31} \approx 4.66 \times 10^{-10}$	$\pm 1.$	semicircles
C_{uc}	16	$2^{-29} \approx 1.86 \times 10^{-9}$	$\pm 6.10 \times 10^{-5}$	radians
e	32	$2^{-33} \approx 1.16 \times 10^{-10}$	.5	dimensionless
C_{us}	16	$2^{-29} \approx 1.86 \times 10^{-9}$	$\pm 6.10 \times 10^{-5}$	radians
$\sqrt{A}$	32	$2^{-19} \approx 1.91 \times 10^{-6}$	8192	meters$^{1/2}$
t_{oe}	16	$2^{4} = 16$	604,784	seconds
Spare	6	--	--	--
C_{ic}	16	$2^{-29} \approx 1.86 \times 10^{-9}$	$\pm 6.10 \times 10^{-5}$	radians
Ω_o	32	$2^{-31} \approx 4.66 \times 10^{-10}$	$\pm 1.$	semicircles
C_{is}	16	$2^{-29} \approx 1.86 \times 10^{-9}$	$\pm 6.10 \times 10^{-5}$	radians
i_o	32	$2^{-31} \approx 4.66 \times 10^{-10}$	$\pm 1.$	semicircles
C_{rc}	16	$2^{-5} = .03125$	± 1024	meters
ω	32	$2^{-31} \approx 4.66 \times 10^{-10}$	$\pm 1.$	semicircles
$\dot{\Omega}$	24	$2^{-43} \approx 1.14 \times 10^{-13}$	$\pm 9.54 \times 10^{-7}$	semicircles/sec
AODE	8	$2^{11} = 2048$	524,288	seconds
Spare	14	--	--	--

* ($\pm$) indicates that the sign bit shall occupy the most significant bit (MSB).

NOTE: All binary numbers will be two's complement.

or

$$e_{T_j} < \frac{.0214 \text{ meters}}{\left.\frac{\partial R}{\partial x_j}\right|_{x^*}} \tag{36}$$

This inequality was used to determine the most significant bit of e_{T_j}. The LSB scale factors of the components of x_T were chosen to be one bit larger. Fractions of bits were dropped. Testing of the algorithm with truncated parameters proved the LSB scale factors to be satisfactory. (See Figure 2.)

VI. DATA BLOCK 3 - THE ALMANAC

Data Block 3 appears in the fifth subframe, appearing every 30 seconds. However, it does not repeat itself every 30 seconds as the other two data blocks do. There are twenty-five subframes of data in Data Block 3, appearing in sequence in the fifth subframe; one every 30 seconds. Thus, each of these twenty-five subframes repeats every 750 seconds.

Data Block 3 is generated by the Control Segment and contains almanacs for up to twenty-four Space Vehicles. The twenty-fifth subframe is a dummy almanac with a Space Vehicle Identification (ID) set to zero. This subframe is present because of a User requirement to have an odd number of almanac subframes to aid in their Data Block 3 data collection. Prior to the availability of twenty-four Space Vehicles, certain Space Vehicle Almanacs will be repeated.

The Space Vehicle Almanacs contain ephemeris representation parameters, clock correction parameters, Space Vehicle ID and Space Vehicle health. The purpose of the almanacs is to provide the Users with less precise Space Vehicle position and clock correction information (relative to Data Blocks 1 and 2 precision) and vehicle health to aid in their direct acquisition of the Space Vehicles' signals. Both the ephemeris representation and clock correction representation in the almanacs are truncated versions of the respective representations in Data Blocks 1 and 2.

Almanac Representation Models

The representation model for the Space Vehicles' clock corrections in the almanacs is identical to that given in Eq. (9) with a_2 assumed to be zero. That is,

$$\Delta t_{SV} = a_0 + a_1 (t - t_{oa}) \tag{37}$$

where t_{oa} is the Data Block 3 reference time in seconds. a_1 is the same as given in Eq. (9) except that it is truncated because less accuracy is required. a_0 can easily be found by integrating Eq. (9) to the almanac reference time t_{oa}. That is, neglecting a_2,

$$a_0 \Big|_{\text{Data Block 3}} = a_0 \Big|_{\text{Data Block 1}} + a_1 \Big|_{\text{Data Block 1}} (t_{oa} - t_{oc}) \tag{38}$$

The representation model for the Space Vehicles' ephemerides in the almanacs is identical to that given in Tables 5 and 6 with certain parameters assumed to be zero. The parameters not assumed zero are the basic seven Keplerian parameters (M_o, t_{oa}, e, $\sqrt{A}$, Ω_0, $i_o + \delta i$, ω) and one perturbation parameter, the rate of change of the right ascension $\dot{\Omega}$. For the almanac, a nominal inclination angle i_o of 60 degrees is defined, and a perturbation δi is used for the sake of saving data bits.

The almanac reference time, t_{oa}, is the multiple of 4096 seconds just prior to 3.5 days after the time that the almanac begins transmission. The almanac will be renewed every six days as a minimum rate. Therefore, the reference time is not ambiguous even though GPS time never spans more than one week. GPS time t will never differ from t_{oa} by more than 3.5 days. The time from epoch t_k of Table 6 should be computed as described in that table, except that t_{oe} is replaced with t_{oa}. However, if the User wishes to extend the use time of the almanac beyond the time span during which it is being transmitted, he must account for crossovers into time spans where these computations of t_k are not valid.

This may be accomplished by computing t_k at the GPS time t_c that the almanac was collected, and storing it as t_{kc}. That is,

$$t_{kc} = t_c - t_{oa} \tag{39}$$

corrected for end of week crossover. The time of year t_{cy} corresponding to t_c should also be computed and stored as

$$t_{cy} = (D_{tc} - 1) \times 86,400 + t_c \bmod 86,400 \tag{40}$$

where D_{tc} is the day of year at the Greenwich Meridian at time t_c. If the GPS time of its use is t_u, the time of year t_{uy} corresponding to t_u is then

$$t_{uy} = (D_{tu} - 1) \times 86,400 + t_u \bmod 86,400 \tag{41}$$

where D_{tu} is the day of year at the Greenwich Meridian at time t_u, corrected for crossovers into new years since the time of collection (i.e., add 365 or 366 for each crossover). The time from epoch t_k at that time is simply

$$t_k = t_{uy} - t_{cy} + t_{kc} \tag{42}$$

which is valid even during the time span during which the almanac is being transmitted. For almanacs that are not collected, but are furnished from an external source, it suffices to define the time of collection t_c as the recording time and the day of year as the recording day of year. This time and day of year will accompany the almanac and will always be within 3.5 days of the reference time t_{oa}.

Almanac Representation Accuracy

For direct P code acquisition in a specified period of time, the desired combined accuracy for the almanacs, including both the ephemeris and the clock corrections, is about 3000 to 6000 meters for up to one week of applicability. However, for direct P code acquisitions in larger periods of time, or for normal C/A code acquisitions, almanacs that are more than one week old, and thus with degraded accuracy, may be used.

The accuracy of this ephemeris representation, including truncation errors, is given in Table 9 for up to five weeks of application. The accuracy of the clock correction parameters is primarily affected by this truncation of the parameters, which may be computed for all time as

$$\mathrm{UERE}_c = \frac{c}{\sqrt{3}}\sqrt{2^{-34} + 2^{-70}\, t_k^{\,2}} \text{ meters} \tag{43}$$

one sigma, where c is the speed of light in meters per second and t_k is as given in Eq. (42).

Table 9. Almanac Ephemeris Representation Performance

Time Past Initial Transmission	Estimated UERE (Meters)
1 day	1,000
1 week	2,500
2 weeks	5,000
3 weeks	10,000
4 weeks	15,000
5 weeks	20,000

Therefore, for one week of applicability, the combined accuracy is about 3200 meters one sigma, and at 5 weeks the combined accuracy is about 24,300 meters one sigma. This is about 828 P code chips or 82.8 C/A code chips, and may be as accurate as, or more accurate than, the User's knowledge of his own position and time, and thus, an excellent aid for faster acquisition.

Data Block 3 Format

Data Block 3 occupies the third through the tenth thirty bit words of the fifth subframes. The number of bits, the scale factor of the least significant bit (LSB), which is the last bit received, the range and the units of the parameters are as specified in Table 10. The last word of the subframe is a six bit spare word and is part of the tenth thirty bit word that has two non-information bearing bits. Thus, it does not link with the TLM word of the next subframe through the parity algorithm.

The scale factors of the LSBs were determined through a sensitivity analysis similar to that described for the Data Block 2 parameters. Equation (34) was changed to reflect the increased tolerance and to reflect only seven parameters.

The reference time t_{oa} has an LSB worth 4096 seconds. This has no impact on accuracy since it is a defined parameter. Its range is the largest multiple of 4096 seconds less than 604,800 seconds.

Space Vehicle Health

The health word in the almanacs provides the Users with a priori information about the applicable Space Vehicle before they attempt to acquire it. The health words are generated by the Control Segment based on its assessment of the vehicle's health. The contents of this word are partly of a preliminary nature and will not be described in this chapter. They are described in detail in Reference 2.

Space Vehicle Identification (ID)

The Space Vehicle ID word specifies the PRN code assignment of the Space Vehicle. There are only 32 such codes assigned to Space Vehicles (1 through 32). The 0th Space Vehicle is a dummy Space Vehicle whose almanac occupies the 25th subframe of the almanacs. With this ID, there are 63 possible IDs requiring six bits of the eight bit word. The other two bits will be used specifying numbers between 0 and 3. These bits will indicate modifications to the standard Navigation Data structure that might occur in the future. At present there are no modifications and zero will be used.

Table 10. Data Block 3 Parameters

Parameter	No. of Bits	Scale Factor (LSB)	Range*	Units
ID	8	1	255	discretes
e	16	$2^{-21} \approx 4.77 \times 10^{-7}$	0.03125	dimensionless
t_{oa}	8	$2^{12} = 4096$	602,112	seconds
δi	16	$2^{-19} \approx 1.91 \times 10^{-6}$	± 0.0625	semicircles
Health	8	1	255	discretes
$\dot{\Omega}$	16	$2^{-38} \approx 3.64 \times 10^{-12}$	$\pm 1.19 \times 10^{-7}$	semicircles/sec
$\sqrt{A}$	24	$2^{-11} \approx 4.88 \times 10^{-4}$	8192	meters$^{1/2}$
Ω_o	24	$2^{-23} \approx 1.19 \times 10^{-7}$	$\pm 1.$	semicircles
ω	24	$2^{-23} \approx 1.19 \times 10^{-7}$	$\pm 1.$	semicircles
M_o	24	$2^{-23} \approx 1.19 \times 10^{-7}$	$\pm 1.$	semicircles
a_0	8	$2^{-17} \approx 7.63 \times 10^{-6}$	$\pm 9.77 \times 10^{-4}$	seconds
a_1	8	$2^{-35} \approx 2.91 \times 10^{-11}$	$\pm 3.73 \times 10^{-9}$	sec/sec
SPARE	6	--	--	--

* ($\pm$) indicates that the sign bit will occupy the most significant bit (MSB).

$i_0 = 0.333333333$ semicircles

NOTE: All binary numbers will be two's complement.

VII. THE MESSAGE BLOCK

The Message Block occupies the third through tenth thirty bit words (including parity) of the fourth subframe. This block provides space for the transmission of twenty-three eight bit ANSC II characters. The remaining eight bits will be non-information bearing.

This Message Block will be generated by the Control Segment. The purpose of the message is to convey alphanumeric information to the Users. It was included in the GPS Navigation Message for future operational applications.

VIII. THE GPS NAVIGATION MESSAGE FRAME FORMAT

The preceding discussions describe five subframes of signal data that make up the GPS Navigation Message, which is called the data frame (1500 bits). The format of the frame was described in part throughout these discussions as the formats for the TLM and HOW words, Data Blocks 1, 2, and 3 and the Message Block. The GPS Navigation Message Frame Format brings these parts together, which is summarized in Figure 3.

This frame repeats itself every 30 seconds, except that Data Block 3 rotates through 25 subframes of data. Periodically (nominally every hour), the data in Data Blocks 1 and 2 are refreshed with data that applies to the new period. Data Block 3 is refreshed by Space Vehicle upload only.

The subframes will always be synchronized to GPS time (within one millsecond). That is, the subframe being transmitted can always be determined from the expression

$$\text{Subframe No.} = \left[\frac{t}{6} \text{ modulo } 5\right] + 1 \tag{44}$$

where [a] indicates greatest integer less than a . Furthermore, the word and bit being transmitted within a subframe can be determined from the expressions

$$\text{Word No.} \quad \left[\frac{t}{0.6} \text{ modulo } 10\right] + 1 \qquad (45)$$

$$\text{Bit No.} = \left[\frac{t}{0.02} \text{ modulo } 300\right] + 1 \qquad (46)$$

IX. FUTURE CONSIDERATIONS

The design, contents and format of the GPS Navigation Message were presented. During the design process, great emphasis was placed on legacy requirements to insure that the design would extend to the operational GPS. Therefore, there is little left for future considerations. There are, however, some parts of the message where changes can be expected in the future. These are as follows:

1) Ionospheric delay correction model
2) Some contents of the TLM and HOW words
3) The health word
4) The non-information bearing bits

Since there are a limited number of Users during Phase I, changes of these parts have essentially no impact if they would occur during Phase I. For instance, the ionospheric delay correction model is only a candidate model being evaluated mostly by the Control Segment. Most Phase I Users who aren't receiving two frequencies will be ignoring the model.

Some contents of the TLM and HOW words are of no interest to the User and are quite Space Vehicle design dependent. They are primarily used by the Control Segment. Also, since the non-information bearing bits may become information bearing bits, additional information can be conveyed via the HOW and TLM. The preamble or system time information should never change.

The health word will be mostly ignored by Users during the early Phase I tests since they will be under complete control of individuals knowing the exact health of the Space Vehicles. It is expected that the contents of this word will evolve during Phase I testing.

The non-information bearing bits may become information bearing bits since future Space Vehicles may generate all of the parity. For now this only impacts the TLM and HOW, the a_0 term in Data Block 1, and the last byte in the Message Block. The reinstatement of these bits will make a_0 more accurate and will allow the Message Block to have another character.

Although it is hoped that all other data will not change, it will not be impossible to change certain parameters during Phase I. All uses of these parameters are programmed in software or firmware. Thus, if changes do occur, only load tapes or the like would have to be changed. Of course, the TLM and HOW words cannot be changed once a Space Vehicle is launched.

X. REFERENCES

1. System Specification for the NAVSTAR Global Positioning System, Phase I, SS-GPS-101B, 15 Apr 74.

2. Space Vehicle Nav System and NTS PRN Navigation Assembly/User System Segment and Monitor Station, Interface Control Document MH08-00002-400, Revision F, 25 July 1977.

3. Van Dierendonck, A. J. and Birnbaum, M., "Time Requirements in the NAVSTAR Global Positioning System (GPS)," Proceedings of the 30th Annual Frequency Control Symposium, June 1976.

4. System/Design Trade Study Report for Global Positioning System Control/User Segment, Trade Study No. 9, General Dynamics Electronics Division Report R-74-034, June 1974.

5. Vessot, R. F. C., "The Relativistic Effects on Clocks and Frequency Standards," Lecture Notes for NBS Seminar on Time and Frequency: Standards, Measurements, Usage, Boulder, Colorado, 22-26 August 1977.

6. Kopitzke, E., "NAVSTAR Global Positioning System Satellite Navigation Ephemeris Algorithm," Magnavox Research Laboratories Report R-5226, 19 Dec 75.

7. Memory and Duty Cycle Requirements for Various GPS User Ephemeris Algorithms, Magnavox Interoffice Communication GPS#76-228, E. Kopitzke to G. Eckley, 29 March 1976.

8. Roy, A. E., The Foundations of Astrodynamics, The Macmillan Co., New York, 1967.

XI. ACKNOWLEDGMENT

The authors acknowledge Major M. Birnbaum (USAF) of the NAVSTAR GPS Joint Program Office for his guidance and coordination with the joint services during the development of this navigation message.

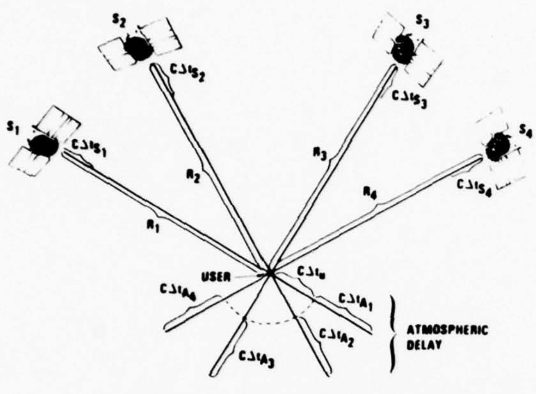


Figure 1. Relationship of Times Between the Space Vehicles and User and their Respective Ranges

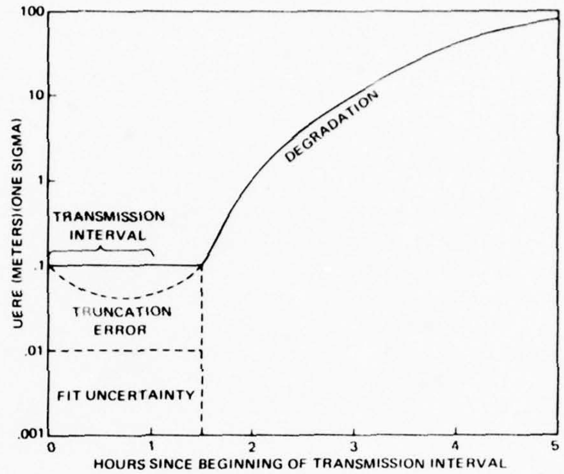


Figure 2. Expected Ephemeris Representation Errors

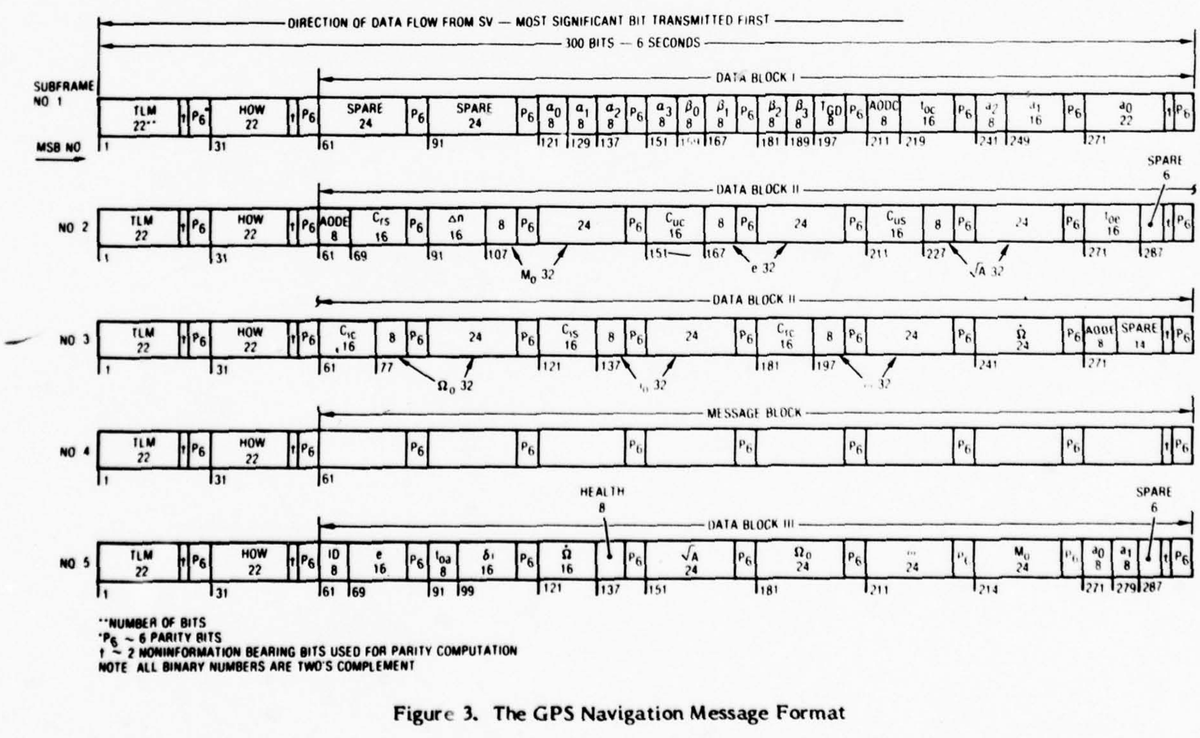


Figure 3. The GPS Navigation Message Format

CLOCKS: EVOLUTION OF FREQUENCY STANDARDS

C. A. Bartholomew*
Space Systems Division
4555 Overlook Ave., S. W.
Naval Research Laboratory
Washington, D.C. 20375

SUMMARY

The clocks currently in use in NAVSTAR GPS Technology Satellites and those planned for future phases of the program are the result of a continuing evolution in clock technology. This evolution is a result of the desire to improve clocks in general and more specifically by the change in character of satellite navigation systems and the need for improvement in clock technology to meet the requirements of such systems. This paper describes the evolution as seen from one program.

INTRODUCTION

Satellite Frequency standards have progressed from quartz crystal oscillators used in the Navy Navigation Satellite System (TRANSIT) and early TIMATION launches to rubidium vapor units first used in Navigation Technology Satellite One (NTS-1), to cesium beam devices currently operating in NTS-2 and hydrogen masers projected for NTS-3.

This evolution has been driven partly by the change in character of satellite navigation systems and partly by a need for improvement in such systems.

TRANSIT operates on a navigation technique suitable for objects having well known velocities. The parameter used is frequency; the satellites operate at low altitudes so the passes last only for 15 minutes or so. The frequency stability of the satellite clock must be such that the change in frequency does not introduce appreciable error in position fix.

The NAVSTAR GPS is designed to give positions continuously in three dimensions. To do so it uses satellites in much higher altitudes than TRANSIT. At these altitudes continuous fixes by means of frequency measurement is impractical so ranging is used instead. Since a further requirement is that the user be passive the ranging measurement is obtained by having all satellites have clocks that are synchronized so the user can make measurements on enough satellites that he can determine the clock synchronization parameters. The problem of clock synchronization without near continuous updating of the satellite clocks has determined the search for better satellite clocks, programming through quartz, rubidium, cesium, and now hydrogen maser standards.

Quartz Standards

TRANSIT used a quartz crystal oscillator designed by APL/JHU as the satellite clock. The principal stability requirements were an overall accuracy of parts in 10^9 with a "short term" stability of a part in 10^{11} for several minutes. Quartz crystal oscillators for laboratory or ground station use were available with stabilities on the order of from a part per 10^{10} to a few parts per 10^{12} per day after an extensive burn in period. The development of this class of oscillator into a flight qualified, low power, highly reliable device with a fractional frequency stability of one part per 10^{12} for averaging times on the order of 10^4 seconds was the first design goal. Such a clock stability would contribute a 10 foot error to the navigator's error budget assuming satellite clock update at 10^4 second intervals.

It should be noted at this time that the current multihundred watt power supplies for spacecraft were not then generally available and the low power and weight and the high reliability of the quartz oscillator made this type a natural choice for the early satellites. The weight, power, and reliability of 1965 era rubidium vapor and cesium beam frequency standards limited their use to ground station functions.

Table one has been prepared to show the sequence of flight experiments to trace the modest experiment that was Timation I launched in 1967 to the multihundred watt satellites that are Navigation Technology Satellites one, two and three of the GPS program.

TIMATION I

TIMATION I was launched into an orbit of opportunity

$$(h = 500 \text{ n.mi}; \quad i = 70^\circ, \quad e = .001)$$

and provided the first demonstration of Navigation based on range measurements from a time synchronized satellite. Figure 1 is a view of the disassembled TIMATION I frequency standard. Figure 2 shows the aging rate for this class of oscillator.

*This work is sponsored by the Naval Electronics Systems Command PME-106.

TIMATION II

TIMATION II was launched into an orbit nearly identical to that of its predecessor. This satellite was designed with a two frequency ranging system to provide ionospheric refraction compensation.

The TIMATION II oscillator included two major improvements. The first was a quartz crystal using thermal-compression-bonded leads, a cold-welded enclosure seal, and a high-temperature bakeout under high vacuum. The second improvement was a triple proportional oven system designed to improve the temperature coefficient by a factor of ten.

A thermal electric temperature-control system was developed for TIMATION II as an experiment to determine whether a critical satellite component could be successfully maintained at a relatively constant temperature above the expected ambient temperature variations. This device, shown in Figure 3, was installed with the oscillator to provide improved thermal control.

Figure 4 is a graph of frequency versus time for the satellite oscillator. The lower curve shows the measured frequency and the times and increments of tuning operations. During the early phase, the aging rate was positive and gradually approaching zero. It was optimistically assumed that the oscillator was recovering from the effects of launch vibration and zero g environment. The low aging rate observed prior to launch (approximately 1 part per 10^{11} per day) and the expected radiation effect could possibly yield a long-term aging rate of a few parts in 10^{12} per day.

The balance of the data shows that this optimism was short lived. The effect of radiation is apparent from the curve, and it can be seen that aging became approximately constant following a period of adjustment due to perturbations caused by orbital injection. It appears that the oscillator has reached a relatively low constant aging rate, but the effect of radiation was not that expected. The upper curve is the normalized frequency-versus-time curve and provides a more graphic view of the results.

A set of improved quartz crystal oscillators were specified for TIMATION III. Figure 5 is a σ - τ plot showing the design goal for these oscillators along with data from three units showing typical performance.

NTS-1

In 1973 the Navy's Timation effort was merged with the Air Force 621B program, with the Air Force named as executive service, to form the NAVSTAR Global Positioning System (GPS) program. NRL's Timation III satellite was redesignated Navigation Technology Satellite One (NTS-1) and launched 14 Jul 74 as part of the NAVSTAR effort.

The launch system selected by the Air Force Space Test Program (STP) for NTS-1 provided sufficient weight margin for the flight of rubidium vapor frequency standards as a first experiment in the newly formed GPS program.

A small, lightweight rubidium vapor standard with low power requirements had just become available commercially from EFRATOM of Munich, Germany. Figure 5 is a photo of the 10 x 10 x 11 centimeter commercial unit which weighed 1.3 kilograms and required a nominal 13 watts of power. Based on the short time schedule and a low budget it was decided that a simple program consisting of only those modifications necessary to make the commercial units flight-worthy would be possible[1]. Feasibility and vibration post-mortem tests were performed on several units to determine the necessary modifications in structure and components. Suitable electrical interfaces were designed to provide tuning of the VCXO and telemetry monitoring.

Six production units were purchased for modification and flight qualification. Of these six modified units, two were selected for flight on the basis of their overall performance under environmental conditions. Figure 7 is a σ - τ plot showing typical pre and post modification performance of the flight candidates. The probable cause of the deterioration is the modification to electronic components particularly in the VCXO area. Figure 9 is a plot of frequency versus temperature in vacuum and in air. The change of temperature coefficient of 7 parts per 10^{13} per ^{o}C to -2 parts per 10^{11} per ^{o}C from air to vacuum is an indication of the lack of optimization of commercial units designed to operate in air. On-orbit performance has been reported[2] on these units however the lack of attitude stabilization on NTS-1 resulted in large temperature variations which ultimately masked any quantitive evaluation of rubidium standard performance. Figure 10 is a coarse plot of frequency versus time which shows variations of frequency caused by temperature variations of the spacecraft. Figure 11 is a long term plot of the frequency of the quartz crystal oscillator which was the principal frequency standard on NTS-1. The general negative trend in frequency is probably due to radiation. Two tuning operations are seen in two years of data.

---

[1] Design and Ground Test of the NTS-1 Frequency Standard, S. A Nichols et all, NRL Report 7904, 5 Sept 75

[2] NTS-1 Quartz and Rubidium Oscillator Frequency Stability Results, Buisson & McCaskill, NRL Report #7932

NTS-2

While NTS-1 rubidiums standards operated in space and showed sufficient capability to be used in other GPS satellites cesium standards offered even greater promise. The promises involved somewhat better long term stability, smaller temperature coefficients and longer life.

Just as the EFRATOM unit became available in time for NTS-1 so also did a cesium unit become available for modification for NTS-2. This unit is based on a cesium tube developed by Frequency & Time Systems, Inc. of Danvers, Mass.

This new cesium tube was particularly attractive in that its dimensions were 7.6 x 7.6 x 30.5 cm with a weight of 4 Kg. A development contract was let to modify the tube to operate through the launch vibration environment and was followed by the development of a prototype frequency standard to be flight qualified and evaluated on the NTS-2 spacecraft. (HP subsequently modified the design of their option 004 tube to meet a tough vibration specification and thereby provided a backup to the FTS prime effort). Figure 12 is a photograph of one of the prototype standards flown on NTS-2. Figure 13 is a σ - τ plot showing the FTS specification along with preflight data taken from the two flight units and the backup. During the first four months of operation of the NTS-2 cesium standard the following milestones have occurred.

Initial frequency measurements precise to one to two parts per 10^{12} have confirmed the calculated offset due to relativity of 4.45 parts per 10^{10}. This offset has been corrected by means of a digital synthesizer in the clock system.

Once the FTS program had completed the critical brassboard demonstration phase a follow-on program for the design of Engineering Development Models (EDM) was initiated. One of these to be included in each of the Navigation Development Satellites (NDS) beginning with NDS 4 as a supplementary frequency standard to the rubidium vapor units. The EDM design included a repackage as shown in Figure 14 to provide a unit with dimensions 12.8 x 19.5 x 38.1 cm weighing 11.3 Kg compared to 25.4 x 30.5 x 40.6 cm weighing 13.6 Kg for the prototype units. A radiation hardening program supported and funded by the Defense Nuclear Agency made a significant contribution to the successful development of a hardened frequency standard meeting GPS requirements. Analysis for this program was by Itelcom Rad Tech (IRT) and FTS. Coordination and much of the test work was by and at NRL.

The latest phase of this work is the current preproduction contract which will provide units for NDS's 7 and 8 in final form.

NTS-3

Cesium clocks require updating to maintain a specified error budget in the satellite system. If the space system is to be made less dependent on the ground system the first need is for an improved clock. The hydrogen maser has been chosen as the best frequency standard offering a significant improvement over the cesium beam. A dual program (Hughes Research Labs and RCA) has been underway for the past year and a half to develop a capability in industry to provide these highly specialized devices in a spacecraft configuration. Figure 15 is a photograph of a test bed maser fabricated under the Hughes program.

A third technology satellite (NTS-3) has been designated to carry advanced development model (ADM) hydrogen maser frequency standards to determine the feasibility of operating the GPS Space Segment independent of ground support for periods such as might be encountered in a "short war".

TABLE I

Technology Satellites

	T–1	T–II	NTS–1	NTS–2	NTS–3
Launch date	5/31/67	9/30/69	7/14/74	6/23/77	10/81
Alt. (N.Mi.)	500	500	7,400	10,980	10,980
Inc. °	70	70	125	63	63
Ecc	.001	.002	.007	.0004	.0004
Wt (Kg)	39	57	295	440	490
Power (W)	6	18	125	400	450
Freq.'s	UHF	VHF/UHF	UHF/L	UHF/L_1/L_2	UHF/L_1/L_2
Clock	Qtz	Qtz	Qtz/Rb	Qtz/C_s	C_s/H–M
$\frac{\Delta f}{f}\left(\frac{PP10^{13}}{day}\right)$	300	100	5–10	1–2	0.1

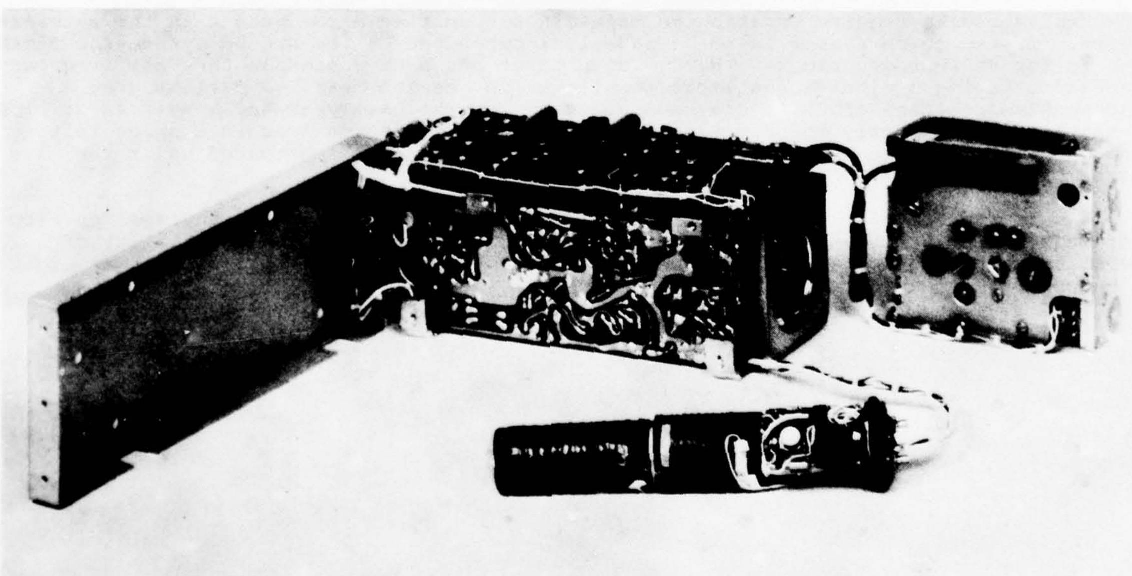


Fig.1 View of disassembled TIMATION I satellite frequency standard

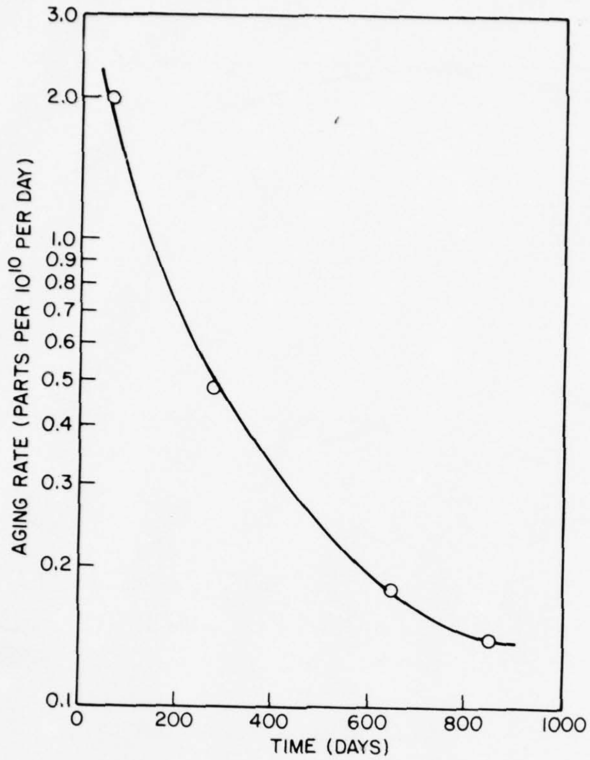


Fig.2 Aging rate of an oscillator of TIMATION I class

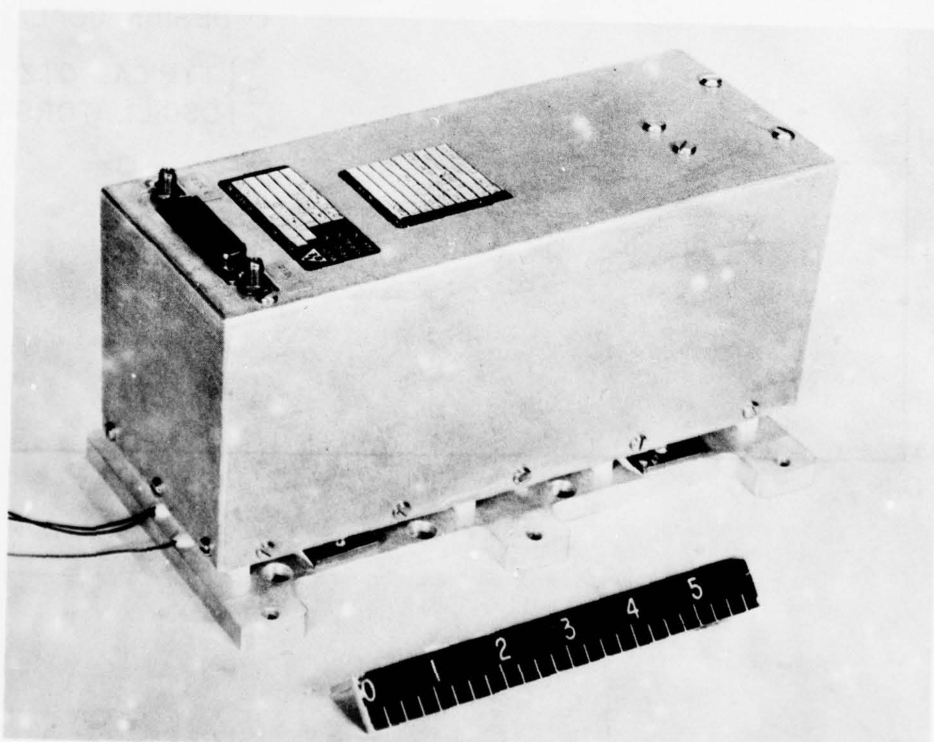


Fig.3 TIMATION II oscillator with thermal electric device (TED) temperature compensation

4. 5 Interface Requirements

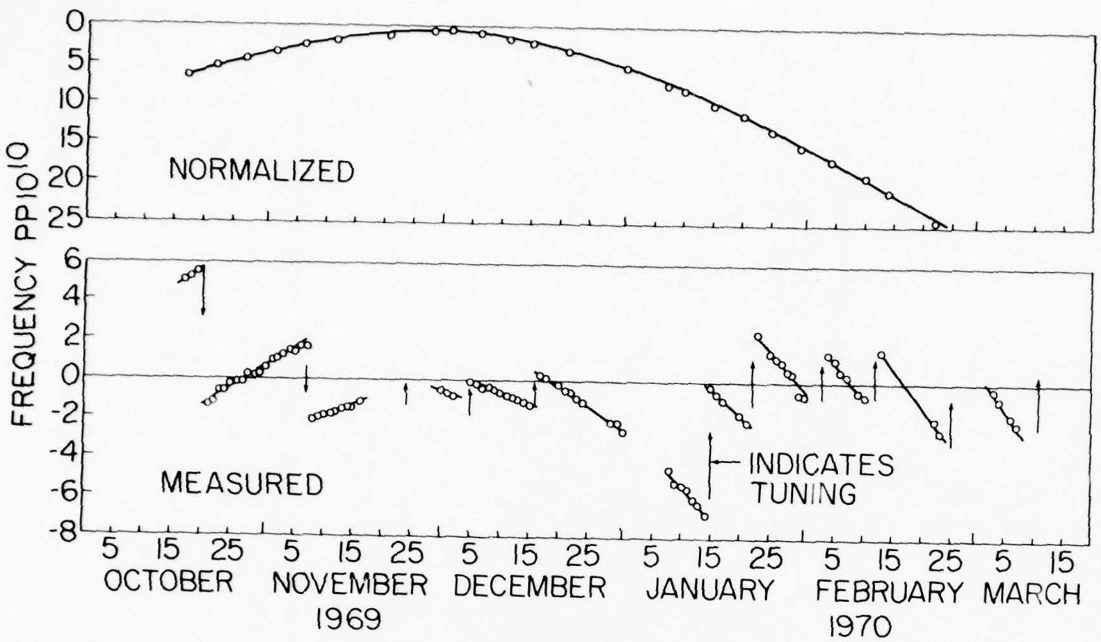


Fig.4 TIMATION II frequency drift

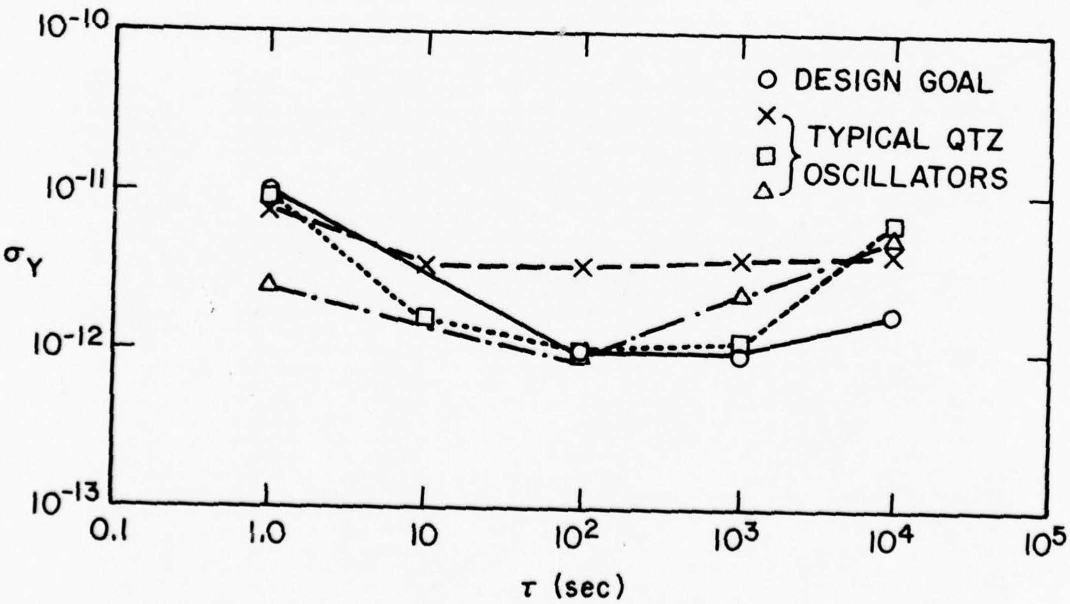


Fig.5 Typical quartz crystal oscillators for spacecraft

where

$$\Delta t_s = a_0 + a_1 (t_T - t_{oc}) + a_2 (t_T - t_{oc})^2 \qquad (8)$$

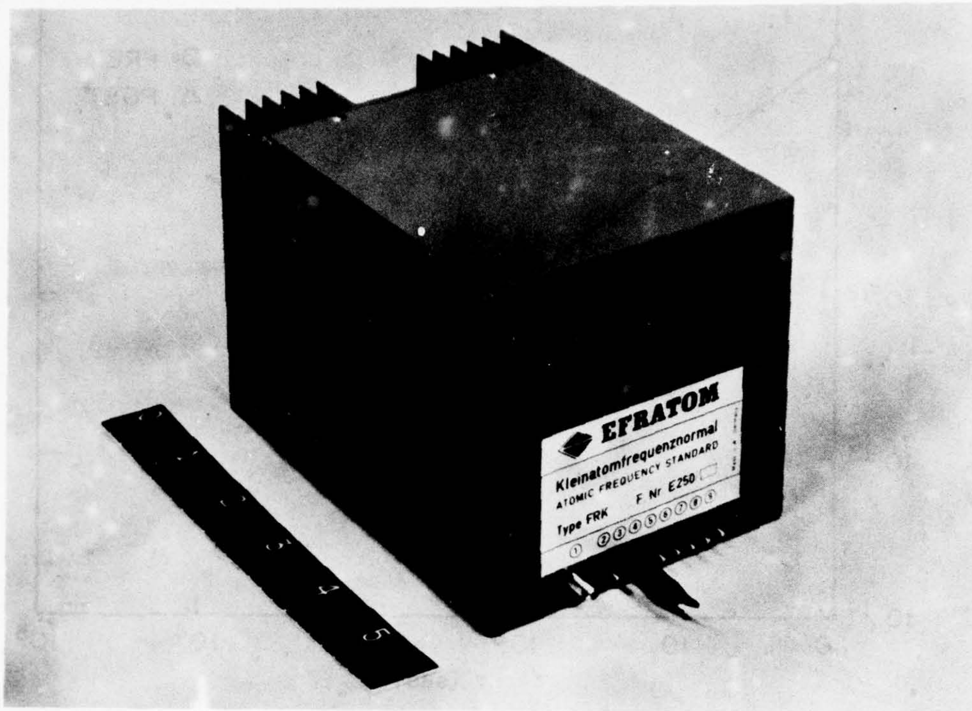


Fig.6 Commercial Rubidium unit by Efratrom

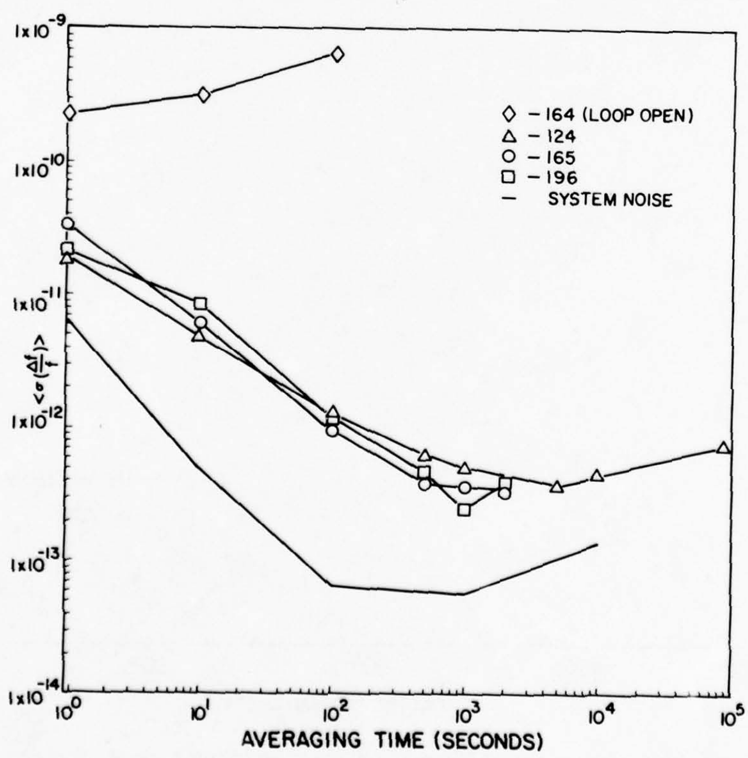


Fig.7 Preflight Rubidium stability data

6.0 CONSIDERATION FOR PHASE III

Relative to GPS time, there are surely numerous i...

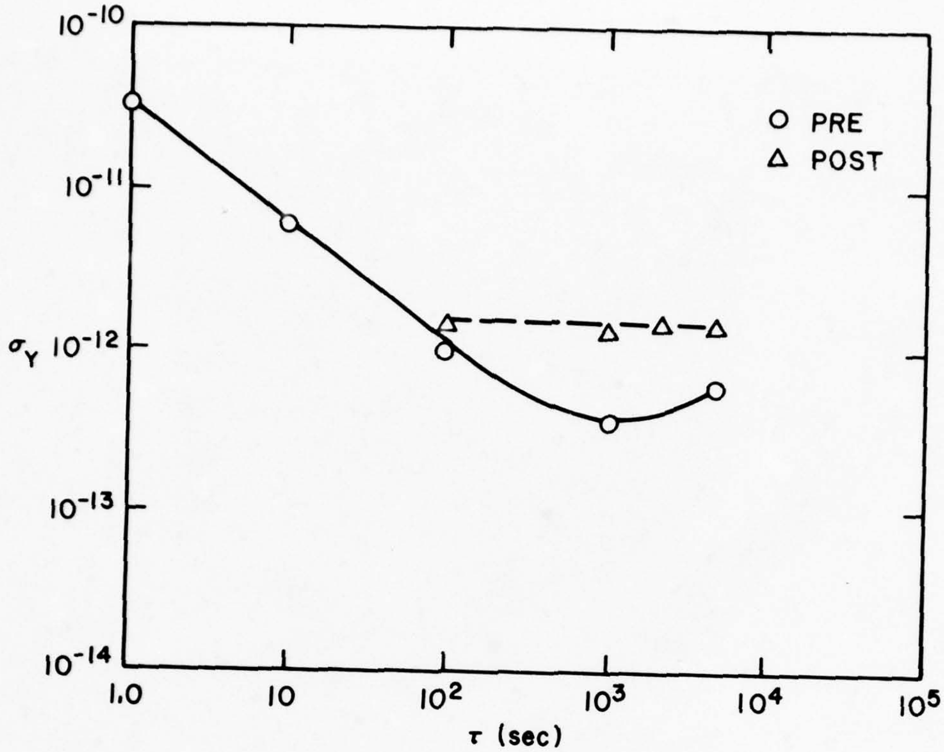


Fig.8 Pre and post modification stability comparisons

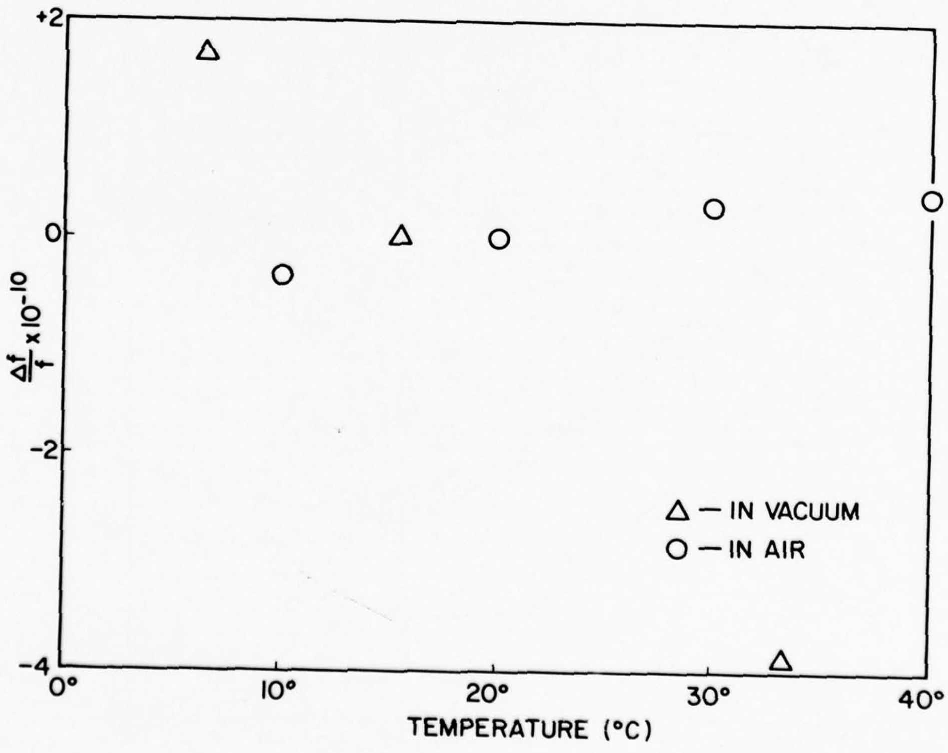


Fig.9 Comparison of temperature performance of Rb units in vacuum and in air

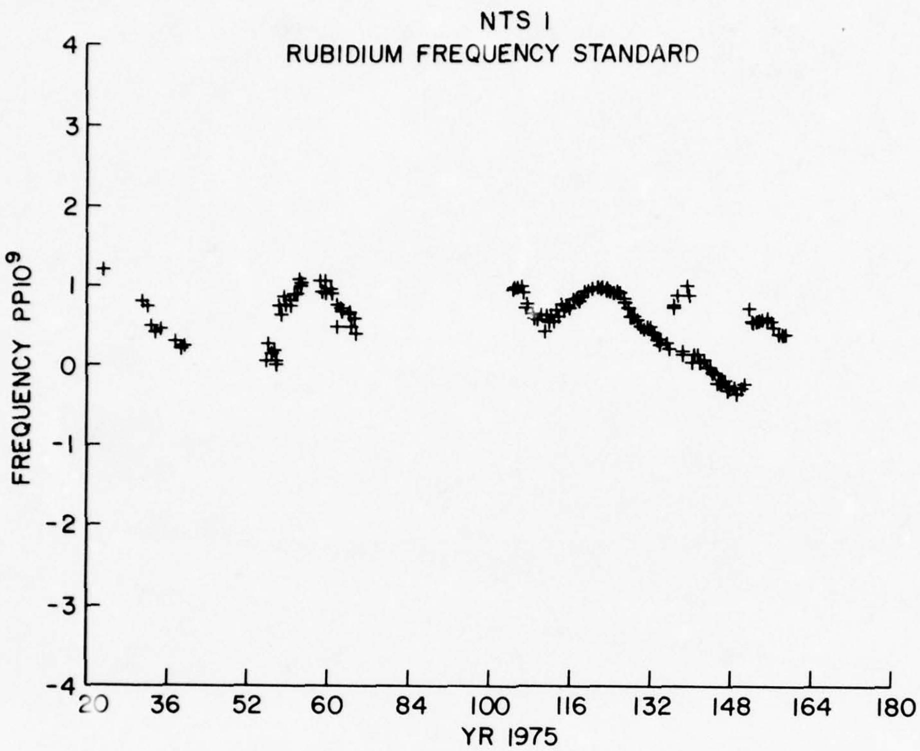


Fig.10 Long term frequency performance of NTS-1 Rb standard

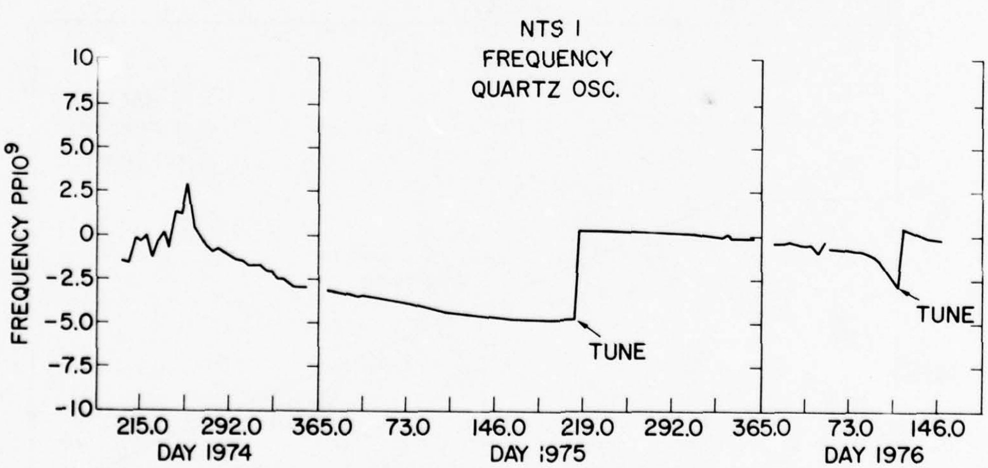


Fig.11 Long term frequency performance of NTS-1 quartz crystal standard

User Data Requirements for Navigation Solution

A nominal three-dimensional navigation fix by a GPS User requires measurements from

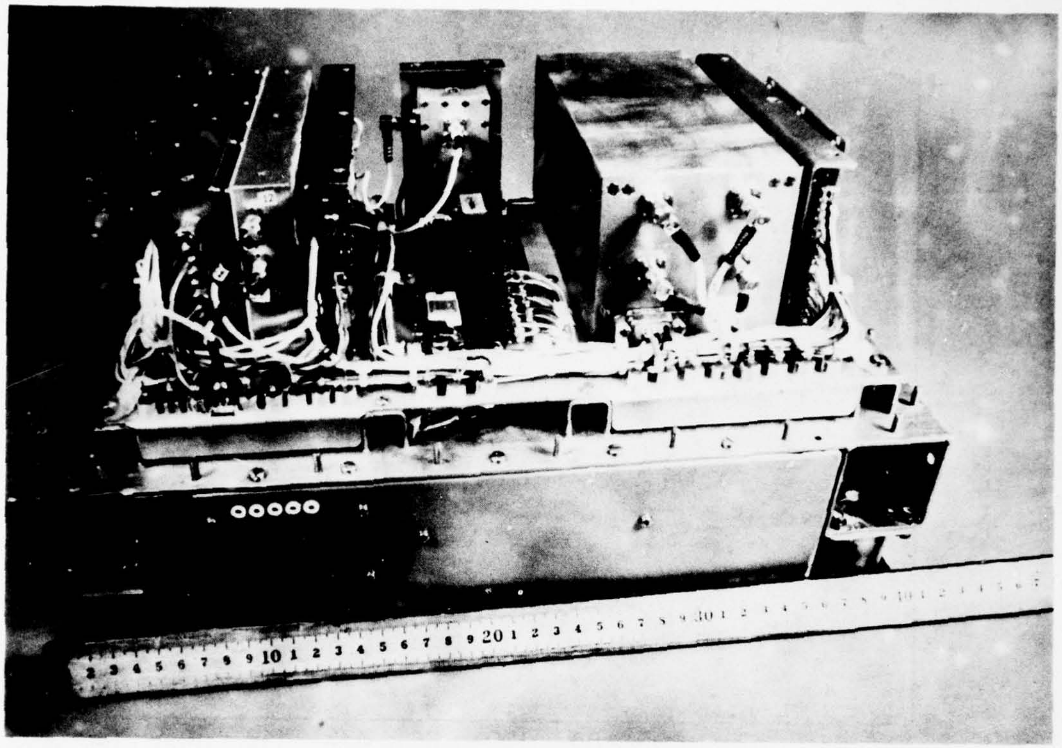


Fig.12 NTS-2 cesium beam frequency standard

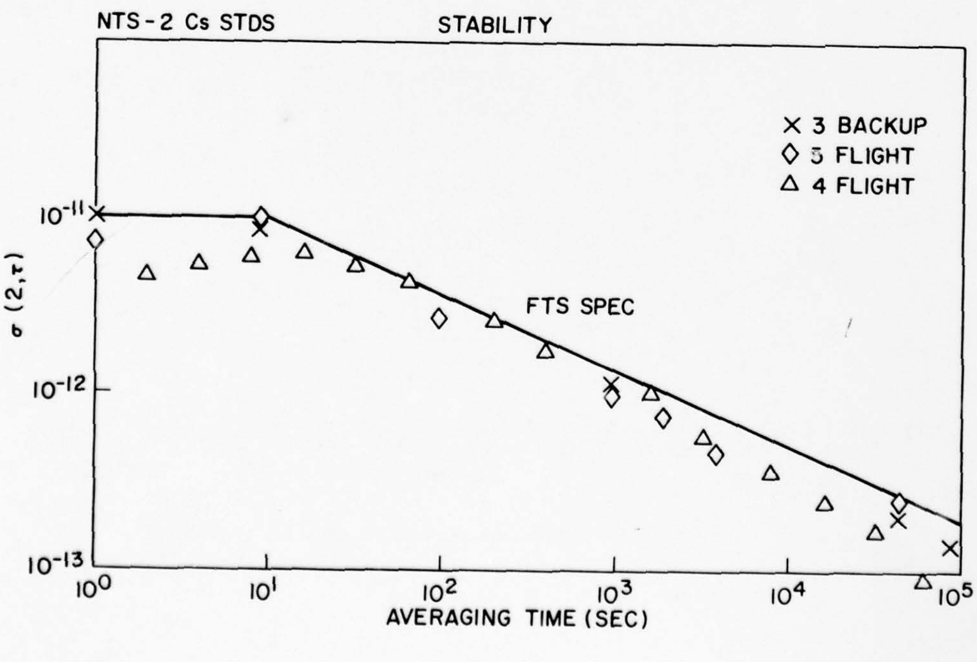


Fig.13 NTS-2 cesium beam freqency standard performance

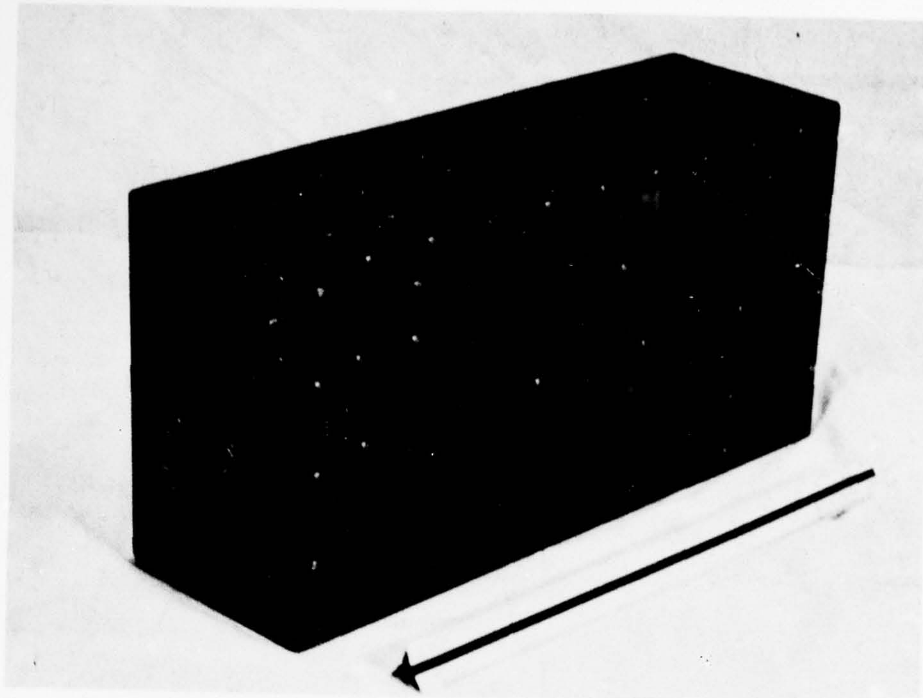


Fig.14 Engineering Development Model cesium beam frequency standard

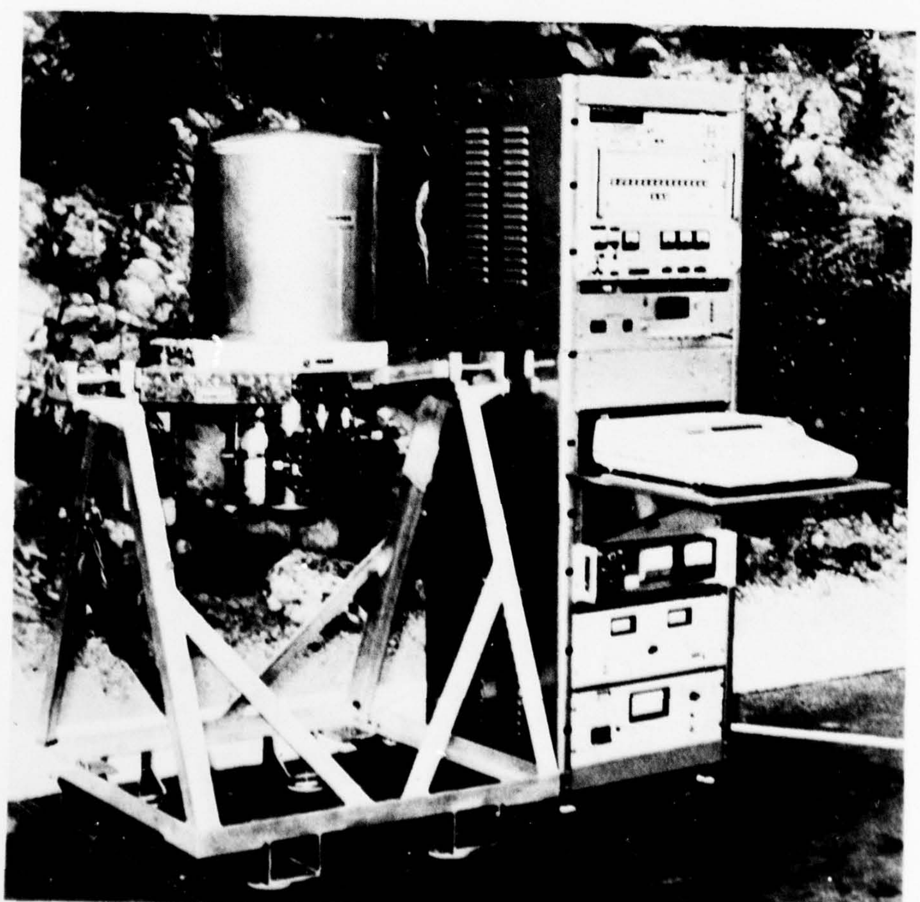


Fig.15 Test bed maser by HRL

GPS TIME

Dr. A. J. Van Dierendonck
AVAND Systems Engineering
6651 Edmonton Avenue
San Diego, California 92122

SUMMARY

GPS navigation is accomplished by one-way time measurements. One nanosecond time error equals 0.984 feet range error. Thus, time control and distribution are essential to GPS. The GPS Control Segment (CS) controls GPS time. The GPS Space Vehicles (SVs) distribute GPS time.

There are three basic features of GPS time derived from the system concept which provide: 1) compatibility with other timing and navigation systems; 2) "fast" User acquisition of GPS "P" signals; 3) continuous User navigation accuracy.

For GPS time to be compatible with other timing and navigation systems places a requirement on the CS to initialize and maintain GPS time to within 100 microseconds of Coordinated Universal Time (UTC). GPS uses external data that are time-tagged with UTC, and other systems will use GPS for time transfer.

For "fast" acquisition of the GPS "P" signal, the User must have the capability of acquiring that signal directly. To do so he is dependent on the a priori knowledge of the SV's times (to within about 10 microseconds) and positions (to within about 10,000 feet). Thus, the CS initializes and controls the SV clocks to within this tolerance, and the SVs clocks are capable of being initialized and controlled to this tolerance. Precise a priori knowledge of time is not required for normal User acquisition of the GPS "C/A" signal.

User navigation accuracy is directly related to his current knowledge of SV clock time. Thus, SV clocks maintain a running time that is predictable to within GPS performance requirements. The CS then predicts their times, generates SV clock update parameters, and uploads these parameters into the SVs.

For providing continuous user navigation, the SVs continuously radiate signals with superimposed navigation parameters to the users. The CS formats these parameters and SV processor control parameters compatible with the SV processor design.

1.0 INTRODUCTION

GPS measurements are corrected transit times of SV generated signals radiated to a receiving User. Since the system is passive, the measurements are meaningful only if the times at which they are transmitted and received are precisely known. The fact is they are not, and the User must correct the apparent transit time with information supplied to him or solved for by him.

In Figure 1, four SVs are shown as required for a three-dimensional navigation solution. This is because the User does not normally know his time - User Time (UT). The offset of this time with respect to GPS time at the time of simultaneous receptions is denoted Δt_u and is common to all four measurements. The GPS reception time is:

$$t_R = UT - \Delta t_u \tag{1}$$

and the GPS transmission time for SV i is:

$$t_{T_i} = t_{Ts_i} - \Delta t_{s_i}, \quad i = 1, \ldots, 4 \tag{2}$$

where t_{Ts_i} is the SV time at time of transmission and Δt_{s_i} is the offset of this time with respect to GPS time. The atmospheric delays from the SV i are Δt_{A_i}. Random variations in the atmosphere, multipath and general transmitter and receiver noises are ignored; c is the speed of light.

The transit time is truly the difference between the GPS transmit time and the GPS receive time. It represents the true slant range except for the propagation delays (the Δt_{A_i}). The User's apparent transit time defines the pseudo-range measurement. The true slant ranges are:

$$R_i = c(t_R - t_{T_i}) - c\Delta t_{A_i}, \qquad i = 1,\ldots,4 \tag{3}$$

while the corresponding pseudo-range measurements are

$$\tilde{R}_i = R_i + c\Delta t_{A_i} + c(\Delta t_u - \Delta t_{s_i}) = c(UT - t_{TS_i}) \tag{4}$$

Obviously, there are the three basic corrections that the User must account for to derive true slant range. The first is the SV clock offset which must be supplied to him; the second is the atmospheric delay, which only he can estimate as it depends on his geometry; the third is his clock offset which only he can determine - the reason for the fourth SV.

Supplying the User with a precise SV clock offset is the joint responsibility of the GPS Control Segment and the GPS Space Vehicle Segment (SVS). This objective and responsibility define the time requirements in GPS. These time requirements in GPS fall in one of three categories:

(1) GPS time synchronization with UTC
(2) SV time and data control for User acquisition
(3) SV time and data control for precise continuous User navigation

The requirements of these categories will be discussed followed by considerations for the Phase III operational GPS. The concept of GPS time discussed here is for the Phase I concept validation of GPS, however, most of them extend to Phase III requirements.

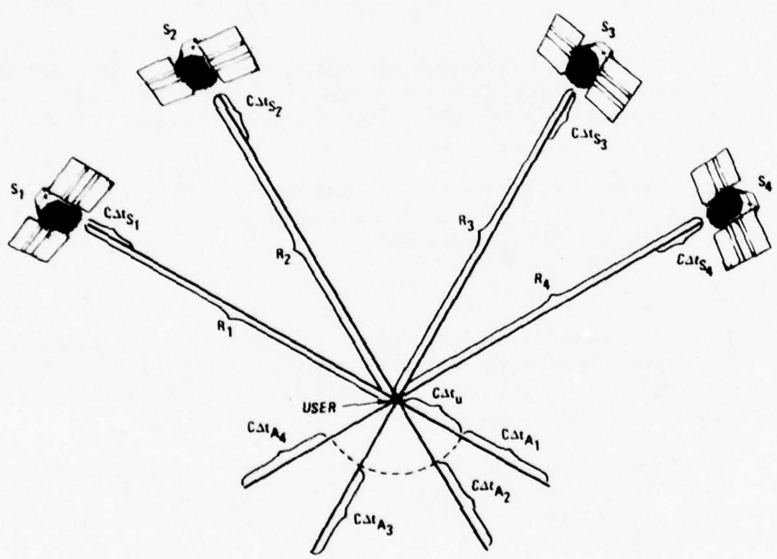


Figure 1. Relationship of Times Between SVs and User and Their Respective Ranges

2.0 GPS TIME SYNCHRONIZATION WITH UTC

2.1 CS Specified Requirement

There is a requirement that the CS reference time shall not deviate from Coordinated Universal Time (UTC) by more than 100 microseconds.[1] The intent of this requirement is to synchronize GPS time with UTC. The purposes of this requirement are both internal to GPS as well as for external systems. Internally, the CS has components which are agencies that have no physical access to GPS time. These agencies - the Air Force Satellite Control Facility (AFSCF) has access to UTC, and the Naval Surface Weapons Center (NSWC) works with data referenced to UTC.

The AFSCF controls the Air Force SVs and provides a backup upload capability for the CS. It also performs the initial tracking of those SVs and provides the initial ephemeris epoch state to the CS for each of those SVs. None of these functions require time synchronization any more accurate than a second or two, but they do indeed require time synchronization.

NSWC periodically provides the CS with a reference ephemeris for each SV. These references are determined using data measured and time-tagged by the CS. Therefore, they are referenced to GPS time. NSWC does, however, as does the CS, use UTC time-tagged Sun and Moon locations in their respective ephemeris determination processes. Again, this does not require an accurate time synchronization, but does require "a" synchronization.

GPS Users, also internal to GPS, benefit by knowing GPS time prior to navigating with GPS. The extent of the benefit depends on how accurate their knowledge is. The maximum benefit is to know it to within 10 to 20 microseconds so as to perform direct "P" code acquisition. This requires that they have a stable frequency standard in order to stay synchronized to GPS time once they have ascertained it. Although synchronizing GPS time to within 100 microseconds to UTC may provide a marginal direct "P" code acquisition, the intent of the requirement is not, at least during the concept validation phase, to provide the capability to initialize Users with UTC.

Certainly a very important use of GPS is for time transfer for initializing or interfacing with external systems. This is where the real requirement for synchronizing GPS time to UTC really lies. For some applications the accuracy of this synchronization certainly has to be much better than the 100 microseconds accuracy that was specified. However, for the Phase I GPS there was no specific application. For later phases of the program, better accuracies are certainly achievable but for now it is set at the 100 microsecond level.

Time transfer appears to be the only concrete reason for this requirement during the initial phase of GPS. Since a 100-microsecond synchronization is not an unreasonable requirement to achieve with the use of a "flying" clock, it has been generally accepted as the requirement.

2.2 Related Requirements

There are four related requirements on the CS that serve to reduce the nuisance factor of maintaining synchronization of GPS time to UTC. These requirements specify the stability, the estimation accuracy and the control of the CS Monitor Station (MS) clocks, which provide the GPS time reference. These requirements are much more stringent than necessary for synchronization to UTC, as they are meant to maintain internal GPS system accuracy. They will be discussed in detail as requirements for SV time and data control for precise continuous User navigation. They do, however, eliminate the need for frequent monitoring and/or reinitializing GPS time with respect to UTC.

2.3 Derived Requirements

Although they are not specifically stated in any GPS system level or segment level specifications, there are requirements imposed on the CS that were derived during the design process. These are the requirements to: 1) initialize an MS clock with an external time source; 2) initialize the Upload Station (ULS) clock; 3) periodically verify the time difference between an MS time and UTC; and 4) publish the time difference between GPS time and UTC.

The reason for the first of these requirements is obvious. It is only necessary to do so at one MS as all other MSs are initialized via signals received from an SV. All MSs are capable of receiving a pulse from a "flying" clock (calibrated to UTC) to initialize a time counter at a time entered into the MS computer. An operator-induced signal prior to the correct pulse enables the counter. This initialized MS is the first "Master" MS as its time is considered perfect (within a known bias). The precise times at the other MSs are estimated through SV tracking and estimation.

The requirement to initialize the ULS (computer) clock is necessary regardless of the difference between GPS time and UTC because it is shut down whenever it is not in use. Time is necessary in the ULS for antenna pointing and for upload verification (within a second or two). The ULS computer or operators never have access to GPS time, but the operator does have access to UTC (via his calibrated watch). As will be discussed later, there can be a significant difference between GPS time and UTC (although known). Since the ULS receives parameters from the Master Control Station (MCS) used in functions of GPS time, either the operator or the computer must correct for the difference.

Of course, GPS time as defined by the "Master" MS and UTC are going to drift apart mostly because frequency standards cannot be perfectly calibrated. Since, the frequency difference won't be known initially, it will be necessary to verify the time difference between GPS time and UTC at later dates. Changes in the time difference may be predicted based on these verifications. To accommodate these verifications the MS clocks provide an output pulse every six seconds. Measuring time between one of these pulses and a pulse from a calibrated "flying" clock accommodates the verification.

To reset GPS time is disruptive to the continuous operation of the GPS system. Because UTC is not a continuous time reference (it has occasional one second jumps to be consistent with the earth's rotation rate), and because GPS time will drift away from UTC, maintaining GPS time to within 100 microseconds of UTC would be disruptive to the GPS system. To have discontinuities in GPS time would cause Users to lose lock from the GPS signals, and would cause extreme difficulties within the

CS to coordinate the discontinuity. Thus, the requirement to merely publish the time difference between GPS time and UTC was derived. This publication is interpreted to meet the intent of the original 100-microsecond requirement.

3.0 SV TIME AND DATA CONTROL FOR USER ACQUISITION

It is essential that the GPS SVs timing be under strict control to enable the Users to acquire their signals. The CS and the SVS are designed to maintain that control.

3.1 GPS Specified Requirements

This time control requirement is reflected in part in the overall GPS System Specification[2] that is referred to by the CS and SVS Segment Specifications.[1,3] The requirement is in the form of a definition of the Navigation Signal Structure. The specification states that the waveform shall be specifically designed to allow system time to be conveniently and directly extracted in terms of standard units of days, hours, minutes and integer multiples and submultiples of the second. It specifies the SVs "P" (precision) signal timing. The Pseudo-Noise (PN) code chipping rate is defined to be 10.23 megabits per second. The code epoch occurs exactly 7 days of elapsed system time after its last epoch by resetting itself routinely at seven day intervals. The measure of this elapsed time shall be the number of X1 (the first of two PN codes Modulo 2 summed together) epochs, termed the "Z" count, which are counted since the PN code epoch. The time between X1 epochs are exactly 1.5 seconds of SV time. Thus, a Z count is worth approximately 1.5 seconds of GPS time (within the accuracy of SV time).

The Z count is transmitted every six seconds, and are contained in the Handover Word (HOW) of the also Modulo 2 summed synchronous data bit stream D, and represents the system time at the start of the next data subframe (a subframe is six seconds long). The Z count value and the time of the PN code epoch are adjustable. Furthermore, all SVs transmit the same Z count within an accuracy dictated by derived requirements.

The C/A (coarse/acquisition) signal code has a chipping rate of 1.023 megabits per second. Its XG code epoch occurs every millisecond. There is a navigation data bit transition every 20 milliseconds, providing data at a rate of 50 bits per second. These bit transitions, the PN code epochs, the XG code epochs, the X1 code epochs and the D data bit stream epochs all occur in synchronization at their respective integral multiple rate. All rates are coherently derived from the frequency standard, and at the same rate on both the L_1 and L_2 frequency carriers, whose frequencies are also derived from that same standard.

The GPS System Specification also specifies the content of the signal data stream D. For the User the signal data has two functions. One is to allow the User to navigate continuously, which is discussed later, and the other is to aid the User to acquire the transmitted signals. Certain timing information in the data is essential for User acquisition:

(1) System time
(2) A preamble for synchronization
(3) Subframe identification
(4) SV clock information for all SVs

The system time is in the form of the Z count in the Handover Word as described earlier. The preamble is a fixed 8-bit word that appears at the start of each subframe (every six seconds) which also coincides with an X1 code epoch and occurs at the Z count transmitted in the previous HOW word. There is also a subframe identification in the HOW word, however, since the data stream epochs occur in synchronization with the PN code epochs, this information is redundant with the Z count.

For normal acquisition on "C/A", this is all the timing information that is required by the User, since once he has obtained synchronization with the data stream and received the Z-count, he is able to "define" pseudo-range, transfer to "P" signal tracking and collect data required for navigation. For efficient direct acquisition on "P", however, the User requires a priori information on the SV whose signal he is trying to acquire. This information may have been prestored by the User, which is only useful if he knows GPS time and his position to within a few thousand meters. To normally know this information, he would normally have been tracking other SV signals. Thus, so that he may obtain the necessary information to directly acquire a new SV's signal, all SV data streams contain information on all SV positions and clocks (almanacs). This information appears in subframe 5 of the data stream, providing almanacs on the SVs on a rotating basis. The information on a given SV appears at least once every 25th frame (made up of 5 subframes), or once every 750 seconds. The clock information required in the almanac is the SV's time offset and drift, which provides the offset to within an accuracy dictated by derived requirements.

Most of these timing functions are inherent in the SV design, however, the CS synchronizes the Z counts on the SVs and provides the almanac contents, both via normal uploading of the SVs.

3.2 SVS Specified Requirements

In addition to the GPS specified requirements, the SVS also has requirements from the SVS Specifications that control the SVs timing to enable User acquisition. These are three requirements on the Clock (oscillator) Assembly accuracy and control. The first is a frequency accuracy. The frequency error of the SV emission shall be less than one part in 10^8 during the life of the SV. However, an error of this magnitude represents a time drift of 864 microseconds a day, and is dominated by derived requirements on the almanac word sizes jointly with the requirements on the frequency of Z count/code phase adjustments.

The second requirement is a requirement for digital tuning. That is, a capability is provided to reset the frequency in steps no larger than one part in 10^{10} over a range of ± 4 parts in 10^9 around the nominal frequency via the TT&C (telemetry) subsystem. The purpose of this requirement is to minimize the SV clock time drift, and thus, minimize the derived requirements stated in the previous paragraph.

The third requirement is the requirement on the code reset. That is, the code phase and Z counter are capable of being reset to the nearest chip (98 nonoseconds) using the CS control link (upload). The necessity for this requirement is to maintain synchronization of timing between SVs and to minimize navigation data word sizes.

3.3 CS Specified Requirements

In addition to the GPS specified requirements, the CS also has two requirements from the CS Specifications that control the SVs timing to enable User acquisition. The first is that the Master Control Station (MCS) generates clock control commands (e.g., digital tuning) when required to maintain system accuracy, and formats these commands for upload to the SV. The formatted Z-count adjustment commands are included in normal CS uploads to the SV. The frequency adjustments, if necessary, will be commanded by the AFSCF via the TT&C command subsystem.

The second requirement is that the MCS be capable of generating an almanac for up to 24 SVs. The accuracy of the almanac is sufficient to allow each user to acquire the GPS navigation signals and get a position fix in a specified time. The almanac data is formatted for upload to each SV and are included in normal CS uploads to the SVs. The almanac clock data is obtained from a prediction of SV clock parameters, which are products of a related requirement on the CS. That requirement is for the generation of clock update parameters, whose accuracy must be much better than required for the almanac. They are generated for User navigation purposes, and will be discussed in detail as requirements for SV time and data control for precise continuous User navigation.

3.4 Derived Requirements

Certain requirements not specified at the GPS system level or segment level were derived for the control of SV time and data to enable User acquisition. These are the requirements on the synchronization of time between the SVs, and thus, a requirement on the size of SV clock correction words in the navigation and almanac data. The derived requirement on synchronizing all SV's time to within one millisecond of GPS time is based on a derived requirement that all SVs transmit the same bit of the navigation data at the same time. Thus, the Users know what data they are collecting based on what time it is, or, for the sequential User (time-sharing one channel among 4 SVs), making it possible to collect data from all visible SVs on a reasonable time-sharing basis.

The time of the PN code epoch is not specified in either of the system level or segment level specifications. It has been defined to be approximately midnight Saturday Night/Sunday Morning Greenwich Mean Time (GMT).

4.0 SV TIME AND DATA CONTROL FOR PRECISE CONTINUOUS USER NAVIGATION

Given that the User is able to acquire and track an SV signal and demodulate the system data, his next task is to navigate to an accuracy dictated by his mission. To do this he needs precise continuous knowledge of the position, velocity and time of up to four SVs in a desirable geometry. The CS and the SVS provide this information to the User. That SV time information is discussed here.

4.1 GPS Specified Requirements

The specification on most User navigation accuracies in the GPS system is expressed indirectly in terms of User Equivalent Range Error (UERE). The navigation accuracy (one sigma) is normally estimated as the UERE times GDOP, where GDOP is the Geometric Dilution of Precision. GDOP is defined as the square root of the trace of the navigation and time error covariance matrix (4 by 4) for the case where the pseudo-range measurement errors have unit variance and are uncorrelated. GDOP depends only on the relative geometry of the User and the SVs. At the same time, UERE is defined to be that uncorrelated portion of the observed range error.

The SV time error is directly a User Equivalent Range Error since it is always along the line-of-sight. It is also separated from the total UERE specification and is specified in the budget for the SV group delay. For Phase I, the specified budget is a UERE of nine feet (≈ 9.146 nanoseconds) (one sigma) that is applicable for two hours after all SVs are updated. For Phase III, the specified budget is a UERE of three feet (≈ 3.049 nanoseconds) (one sigma) that is applicable for twenty-four hours after all SVs are updated. This SV group delay is defined as the summation of delay uncertainty due to effects in the SV such as unmodeled clock drift and uncalibrated delay in signal equipment.

This UERE specification impacts the CS and SVS jointly. It is necessary for the SVS to provide predictable clock time offsets to within the SV group delay specification, allowing for reasonable CS prediction errors. Likewise, there is a necessity for the CS to predict the SV clock time offsets to within the SV group delay specification, allowing for unpredictable random SV clock errors. There is no clear division of the SV group delay error budget between the SVS and the CS, primarily because the stability of the SV group delay affects the CS's ability to predict it.

The GPS System Specification also specifies the content of Subframe 1 of the SV signal data stream, which the CS provides. Subframe 1 consists of TLM and HOW words (provided by the SVS) and Data Block 1 (provided by the CS). Data Block 1 is generated by the CS and contains the frequency standard corrections. Corrections for relativistic phenomena are included within this data.

The purpose of Data Block 1 is to provide the User with SV clock correction information to correct for SV clock offsets to within the specified UERE (one sigma). In addition, because the SV clocks are at a different gravitational potential and are traveling faster than clocks on the earth, their times must be corrected for general relativistic effects. Although these effects wouldn't be classified as SV group delay, they are included in the clock correction information.

4.2 SVS Specified Requirements

In addition to the GPS specified requirements, the SVS also has requirements from the SVS Specifications that control the SVs timing for precise User Navigation. These are requirements on equipment group delay and on timing stability and clock drift. The group delay from clock to radiated output is calibrated prior to launch. The effective group delay uncertainty of the PRN signal is less than 1.5 nanoseconds (one sigma) during normal operations, and less than 2.5 nanoseconds (one sigma) during SV eclipse operations. Equipment group delays are equivalent to time offsets. Therefore, uncertainties which are biases, or very long time-constant variations, are of no consequence because they appear to be a clock time offset or drift. However, random or short time-constant equipment group delays affect the CS's ability to predict the SV clock drift. The purpose of this requirement is to minimize that effect.

Regarding timing stability, for all on-orbit conditions, the frequency stability is within the requirements specified in Figure 2. This figure represents the square root of Allan 2-sample variance[4] specification on clock stability. This translates to a clock time uncertainty versus time that is presented in Figure 3. In this figure, time drift representing the specified stability of Figure 2 is compared to that obtained in testing various frequency standards, some of which are standards being launched in GPS SVs. * These clock time uncertainties were determined by converting Allan 2-sample variance to continuous time variances. They do not include prediction errors of the CS.

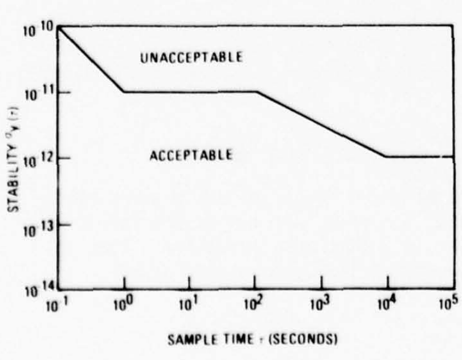


Figure 2. Clock Frequency Stability

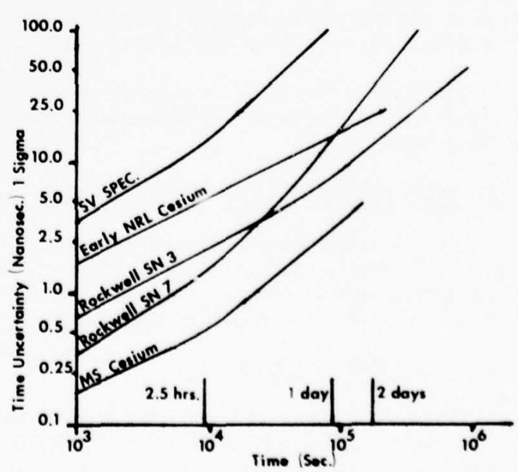


Figure 3. Estimated Clock Phase Prediction Error for Tested Standards

*Provided by Rockwell International Space Division

4.3 CS Specified Requirements

The CS provides the content of Data Block 1 which the User applies to correct for the SV time offset. The MCS processes retrieved data to generate a clock update (i.e., clock bias, clock frequency offset, and relativity effect) for each SV (up to 24) accurate to the level specified in the GPS Error Budget. The MCS generates these clock update parameters relative to the clock that is designated by the operator to be the "Master" clock. The MCS formats all SV update data consistent with the GPS signal structure, generates the SV data files and transmits them to the ULS. The ULS is the primary means for uploading the GPS navigation subsystem with valid navigation data.

4.4 Related Requirements

In order for the CS to predict and generate SV clock update parameters to the accuracy level specified in the GPS Error Budget, it has the following design features to maintain its internal timing:

(1) Stable and accurate frequency standards for the MS clocks
(2) Backup power for the MS clocks to provide continuity of time
(3) The operator may choose which clock is the "Master" clock without destroying the continuity of time
(4) The MCS corrects pseudo-range measurements for SV clock relativistic effects
(5) The MCS performs an optimal estimate of all CS clocks (part of the overall estimation process)
(6) The MCS has an on-line system performance evaluation function that monitors the characteristics of all GPS clocks

Each MS utilizes a frequency standard for deriving its reference timing functions. The standard has a stability (the square root of Allan variance) of 1 part in 10^{11} for one second averaging and 1 part in 10^{13} for 10^4 seconds averaging. The standard does not deviate from its assigned frequency (using averaging periods of 100 seconds or longer) by more than 5 parts in 10^{12} for up to one year after calibration. High performance Cesium beam frequency standards are used in the MSs. The purpose of this high performance stability is to enable an estimation of SV clock parameters without influence from the characteristics of the MS clocks.

With respect to power outage, the MS components, e.g., time, timing, and frequency references, which require continuous power have backup power supplies that will provide 2 hours of operation in event of power outage. Besides reducing the nuisance factor, the backup power provides for a continuity of time in the MSs so that re-initialization and estimation of MS time will only be required after long-term power failures. Re-initialization and estimation puts the SVs in the role of providing the GPS time reference instead of the MSs, thus degrading the capability to estimate and predict SV time.

Because power outages and/or other equipment failures are bound to occur, the MCS provides the operator with the means to select which CS clock is used as the "Master" clock. If the "Master MS loses its time reference, or if its time reference is in a degraded condition, another MS is designated the "Master" MS with a negligible shift in GPS time (expected to be no worse than 2 to 3 nanoseconds with a step change in frequency of no worse than 1 part in 10^{12}). This requirement ensures that there will always be a CS time reference of a quality required to estimate and predict SV time.

Secular relativistic effects on SV clocks, although significant, are not important to the estimation of SV clock parameters or SV ephemerides because they are not distinguishable from deterministic SV clock frequency offsets. These apparent offsets are common to both the CS and the User. Thus, they can be absorbed in the predicted SV clock parameters. However, there are also periodic relativistic effects on the SV clocks that are impossible to absorb in reasonable clock estimation models. Therefore, it is important that these periodic effects be accounted for in CS measurement processing as well as for SV clock parameter updates. Thus, the MCS corrects the ranging data for relativistic effects.

To predict and generate SV clock parameters, the MCS performs an optimal estimation of all CS clock update parameters. Since the MSs are providing the time references for estimating the SV clock parameters, the relative difference between each non-"Master" MS clock and the "Master" MS clock is also estimated.

The system performance evaluation function, although not critical to the estimation of SV clock parameters, provides for a real time evaluation of the estimation and prediction process as well as the health and performance of the various system clocks. Specifically, the MCS provides performance evaluation of: (1) residuals between the latest observed SV clock state and the state predicted by clock parameters in the navigation data frame (an evaluation of the difference between what the User is receiving and what the CS is currently estimating); (2) residuals between the estimate of the SV clock state at the latest epoch and the state which was predicted for that epoch 1, 4, 12, and 24 hours before (an evaluation of the stability of the SV clock); and (3) the uncorrected, synchronization error between the GPS time standard and all MS and SV clocks (an at a glance evaluation of all the system clocks as compared to GPS time to monitor the performance of the MS clocks and to ensure the required synchronization between all clocks and GPS time).

Fig.3 TIMATION II oscillator with thermal electric device (TED) temperature compensation

4.5 Interface Requirements

The requirements affecting the designs in three segments - the CS, the SVS and the User Segment (US) were derived jointly by the three segments. The vehicle for deriving these requirements is an Interface Control Working Group made up of members of the three segments. The resulting document is an Interface Control Document (ICD)[5] referenced in the GPS system and all segment level specifications. This document imposes derived requirements on all three segments. The requirements regarding SV time and data control for precise continuous User navigation are:

(1) SV group delay differential between L_1 and L_2 "P" signals
(2) L_1-L_2 correction for single frequency Users
(3) The format of Data Block 1
(4) The User algorithms for SV clock corrections

The first of these requirements is only important to the two-frequency User if the group delay differential is random or time varying with a short time constant, for the same reasons discussed earlier for the requirements on equipment group delay. The group delay differential between the radiated L_1 and L_2 "P" signals consists of random plus bias components. The mean delay differential is defined as the bias offset in the (L_1-L_2) P delay differential. For a given Navigation Subsystem redundancy configuration, the mean delay does not exceed 15.0 nanoseconds. The random variations about the mean do not exceed 1.5 nanoseconds (one sigma). Since both the CS and the two-frequency User use the L_1-L_2 difference for ionospheric corrections, any differential delay is multiplied by a factor of 1.5625. Thus the 15 nanoseconds results in a 23.44 nanosecond SV clock offset. Because this would be normally absorbed in the SV clock offset estimate and prediction, it has no effect on the two-frequency User. It does have an effect on the single-frequency User, since the observed clock offset to him does not include this delay.

As a safeguard against an appreciable group delay differential between the SV L_1 and L_2 "P" signals, an L_1-L_2 correction is provided for the single-frequency User. An L_1-L_2 correction term, T_{GD}, is an estimated correction term to account for SV group delay differential between L_1 and L_2. This correction is only for the benefit of L_1 only or L_2 only Users because SV clock corrections are based on two frequency corrections. This User corrects SV PRN code phase time, t_{Ts_i}, with the equation

$$t_{Ts_i} = t_{Ts_i} - T_{GD} \tag{5}$$

The CS estimates T_{GD} during periods when the ionospheric delay is at a minimum.

The CS provides the content of Data Block 1 as described in the ICD. In addition to T_{GD} and eight ionospheric model coefficients, Data Block 1 contains the SV clock correction parameters. These parameters are three polynomial coefficients a_o, a_1 and a_2, a reference GPS time since weekly epoch, t_{oc}, and an age of data (clock), AODC. The polynomial describes the SV PRN code phase (clock) offset, Δt_{s_i}, with respect to GPS time, t_{T_i}, at the time of data transmission. These coefficients describe the offset for the interval of time (one hour as a minimum) in which the parameters are transmitted. The polynomial also describes the offset for an additional one-half hour (i.e., one-half hour subsequent to the beginning of transmission of the next set of coefficients) to allow time for the User to receive the message for the new interval of time. The age of data word (AODC) provides the User with a confidence level in the SV clock correction. AODC represents the time difference (age) between the Data Block 1 reference time (t_{oc}) and the time of the last measurement update (t_L) used to estimate the correction parameters. That is,

$$AODC = t_{oc} - t_L \tag{6}$$

The CS also formats the Data Block 1. Of importance in these discussions is the range and scale factor of the SV clock correction parameters and the L_1-L_2 correction. Table 1 presents the format of these parameters. The scale factors determine the accuracy. They were selected to provide correction accuracy on the order of a nanosecond. The ranges were dictated primarily by the worst case SV clock drift characteristics, the desire to minimize the number of SV clock control commands, and the requirement to synchronize the PN code epochs of the various SVs. The range of a_o (976.6 microseconds) indicates the accuracy of that synchronization.

For continuous SV clock correction the User corrects the time received from the SV with the equation (in seconds)

$$t_{T_i} = t_{Ts_i} - \Delta t_{s_i} \tag{7}$$

where

$$\Delta t_{s_i} = a_o + a_1 (t_{T_i} - t_{oc}) + a_2 (t_{T_i} - t_{oc})^2 \tag{8}$$

Note that equations (7) and (8) as written are coupled. While these coefficients a_o, a_1, and a_2 are generated by using GPS time as indicated in equation (8), the sensitivity of Δt_{s_i} to t_{T_i} is negligible. This negligible sensitivity allows the User to approximate t_{T_i} by t_{Ts_i} in equation (8). The parameters a_o, a_1 and a_2 include all general relativistic effects on the SV clock.

Table 1. Data Block 1 Parameters

Parameter	No. of Bits	Scale Factor (LSB)	Range*	Units
T_{GD}	8	$2^{-31} \sim 4.66 \times 10^{-10}$	$\pm 2^{-24} \sim \pm 5.96 \times 10^{-8}$	Sec
AODC	8	$2^{11} = 2048$	$2^{19} = 524288$	Sec
t_{oc}	16	$2^4 = 16$	604, 784	Sec
a_2	8	$2^{-55} \sim 2.78 \times 10^{-17}$	$\pm 2^{-48} \sim \pm 3.553 \times 10^{-15}$	Sec/sec^2
a_1	16	$2^{-43} \sim 1.14 \times 10^{-13}$	$\pm 2^{-28} \sim \pm 3.725 \times 10^{-9}$	Sec/sec
a_o	22	$2^{-31} \sim 4.656 \times 10^{-10}$	$\pm 2^{-10} \sim \pm 9.766 \times 10^{-4}$	Sec

*(±) indicates that sign bit occupies the most significant bit (MSB)

NOTE: All binary numbers are two's complement.

5.0 SUMMARY OF GPS TIME SYNCHRONIZATION

The time synchronization between GPS time and UTC, the MS clocks, the SV clocks, and the User clocks is summarized in Figure 4. This diagram illustrates the maximum clock offsets between the various clocks and GPS time and an expected or required Phase I tolerance on the estimate of those offsets. The MS and some User maximum offsets (3.1 sec) are dictated by the size of the pseudo-range register in their respective receivers. Normally, this offset will never be more than a few milliseconds. The User tolerance was obtained by multiplying a GPS specification specified UERE of 18 to 24 ft times an expected TDOP of 2.5. (TDOP is the Time Dilution of Precision.)

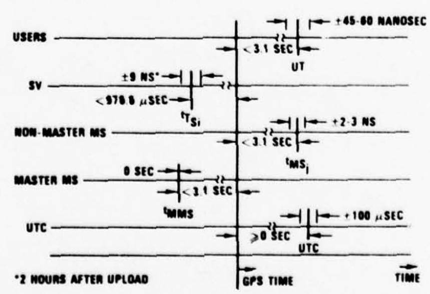


Figure 4. GPS Time Synchronization for Phase I

AVERAGING TIME (SECONDS)

Fig.7 Preflight Rubidium stability data

6.0 CONSIDERATION FOR PHASE III

Relative to GPS time, there are surely numerous improvements that can be made over the Phase I system. Certainly the most important will be the evolution of better frequency standards, especially those qualified for space vehicles. Another very important improvement could be in the ability to estimate and predict their behavior in the space vehicle environment. The concept validation will provide increased knowledge in both of these areas.

Another important consideration for time in the operational GPS is its relationship to the Universal Coordinated Time (UTC). Without very much difficulty, GPS time could be synchronized to UTC to within a few nanoseconds - at least to a known difference. This could be easily realized by placing a GPS User Set in coincidence with a UTC time reference and continually monitoring the time difference between GPS time and UTC. In fact, a GPS User Set or an MS could be slaved to a UTC standard. A normal User set will solve for the difference in time.

The fact that there will be a difference between GPS time and UTC does not pose a problem to the time transfer Users, for in Phase III, that time difference can be transmitted via the navigation data stream as part of Data Block 1.

7.0 ACKNOWLEDGEMENT

The author wishes to express appreciation for the guidance and the DoD community concurrence provided by Major M. Birnbaum of the GPS Joint Program Office during the development of the concept of GPS time.

8.0 REFERENCES

1. System Segment Specification for the Control Segment of the NAVSTAR Global Positioning System, Phase I, SS-CS-101B, 30 Sep 74.

2. System Specification for the NAVSTAR Global Positioning System, Phase I, SS-GPS-101B, 15 Apr 74.

3. System Segment Specification for the Space Vehicle System Segment of the NAVSTAR Global Positioning System, Phase I, SS-SVS-101A, 20 Jun 74.

4. Time and Frequency: Theory and Fundamentals; Byron E. Blair, Editor, National Bureau of Standards Monograph 140, May 1974.

5. Space Vehicle Nav System and NTS PRN Navigation Assembly/User System Segment and Monitor Station, Interface Control Document MH08-00002-400, Revision F, 25 July 77.

TEMPERATURE (°C)

Fig.9 Comparison of temperature performance of Rb units in vacuum and in air

MASTER CONTROL STATION

M. J. Hurley
General Dynamics
Electronics Division
P.O. Box 81127
San Diego, CA 92138
U.S.A.

J. L. Kramer
General Dynamics
Western Data Service Center
P.O. Box 80847
San Diego, CA 92138
U.S.A.

D. D. Thornburg
General Dynamics
Electronics Division
P.O. Box 81127
San Diego, CA 92138
U.S.A.

SUMMARY

Phase I of the GPS navigation satellite effort is to evaluate the performance of User receiving equipment. Phase I is composed of six satellites, a ground Control Segment, and User equipment. The Control Segment generates the data required by the User to obtain a navigation solution, uploads this data into the satellite processor for transmission to the User, and collects the satellite ranging data required to determine the satellite ephemeris and clock performance parameters. The Control Segment software mechanization to perform these functions is a file-based, multi-tasked architecture. This architecture and its legacy to future phases of GPS are described.

INTRODUCTION

GPS Phase I

Phase I of the GPS navigation satellite effort is to evaluate the performance of User receiving equipment on predetermined test ranges. To support this testing, six satellites will be launched to provide a pilot space configuration. The satellite orbits and spatial locations are such that a maximum test period is provided over Yuma Proving Ground (YPG). This is the primary test area for GPS User equipment evaluation.

The Phase I GPS is composed of six satellites, the Control Segment, and User equipment. A representation of the system is shown in Figure 1. The satellites provide highly-stable time-based spread spectrum signal and navigation data to the User. The ground Control Segment (CS) tracks the satellites to determine their ephemerides and atomic clock errors. The CS then predicts the ephemeris and atomic clock model parameters for each satellite. These predictions are reformatted into User navigation data, and uploaded into the navigation processor of the satellite. The User navigation data is then transmitted by the satellite to the User. The User equipment demodulates this data from the spread spectrum signal, and utilizes it in the computation of User position. The primary function of the Control Segment is to generate precise navigation data for the User.

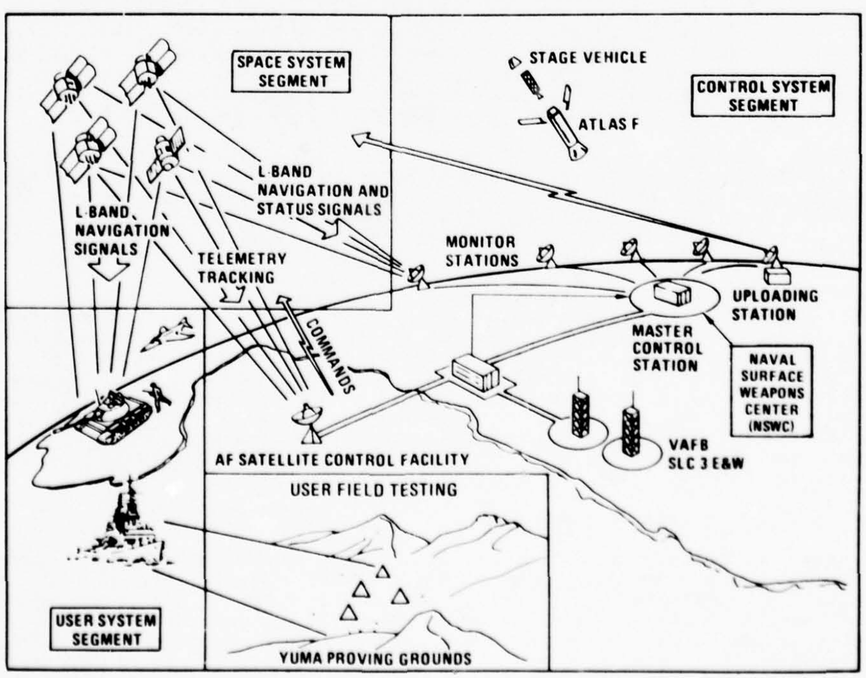


Figure 1. GPS Phase I Overview

User Data Requirements for Navigation Solution

A nominal three-dimensional navigation fix by a GPS User requires measurements from four GPS satellites. Four satellites are required because precise GPS time is assumed unknown to the User. The measurements obtained by the User are referred to as pseudo-ranges. Pseudo-ranges represent the true slant ranges (between User and satellites) except for propagation delays and time biases. The pseudo-range is:

$$\tilde{R}_i = R_i + c\Delta t_{A_i} + c(\Delta t_u - \Delta t_{s_i}),\ i = 1,\ldots,4 \tag{1}$$

where R_i is the true slant range, c is the speed of light, Δt_{A_i} is signal propagation delay, Δt_u represents the time offset between the User and GPS time, and Δt_{s_i} is the satellite atomic clock time offset from GPS times.

As stated previously, the User must solve for four unknowns. They are his earth-fixed, earth-centered coordinates X, Y, and Z and time offset Δt_u. Equation (1) can be written in terms of these unknowns as:

$$\tilde{R}_i = \sqrt{\left(X_{s_i} - X\right)^2 + \left(Y_{s_i} - Y\right)^2 + \left(Z_{s_i} - Z\right)^2} + c\Delta t_{A_i} + c\left(\Delta t_u - \Delta t_{s_i}\right),\ i = 1,\ldots,4 \tag{2}$$

Therefore, the Control Segment must provide a minimum of X_{s_i}, Y_{s_i}, Z_{s_i}, and Δt_{s_i} to the User. In addition, the Control Segment also provides data to permit the User to estimate ionospheric delay, receive special messages, and determine all satellite orbits. Ionospheric delay data permits the User to estimate propagation delay if he is not capable of receiving two frequencies and performing his own delay computation. The purpose of special messages is self-evident. The availability of the almanac of all satellite orbits enables the User to determine satellites in view, aid in satellite selection, and simplify signal acquisition. Because of the importance of satellite position and time offset, this paper will consider the Control Segment role in providing these variables.

CONTROL SEGMENT

The Control Segment consists of four Monitor Stations (MS), an Upload Station (ULS), and a Master Control Station (MCS). The Monitor Stations are located at Hawaii; Elmendorf AFB, Alaska; Guam; and Vandenberg AFB, California. The remote Monitor Stations have been discussed in a previous paragraph and will only be summarized here. The MSs are unmanned data-collection centers under direct control of the MCS. Each MS consists of a four-channel User-type receiver, environmental data sensors, an atomic frequency standard, and a computer processor. The receiver measures the pseudo-range and delta pseudo-range (integrated doppler) of the satellite spread-spectrum signal with respect to the atomic standard. It also detects the navigation data on the spread-spectrum signal. The environmental sensors collect local meteorological data for later tropospheric signal delay corrections at the MCS. The computer processor controls all data collection at the MS, and provides the data interface with the MCS. All data obtained by the MS is buffered at the MS and then relayed upon request to the MCS for processing.

The Upload Station, located at Vandenberg AFB, provides the interface between the Control Segment and the satellites. The ULS has been discussed previously and will only be summarized here. The ULS's function is to utilize an S-band command-and-control uplink to upload data into a satellite navigation processor. This upload data can be User navigation data, requests for processor diagnostics, or commands to change the satellite time provided to the User.

The MCS is also located at Vandenberg AFB, and completely controls the operation of the Control Segment. It performs the computations necessary to determine satellite ephemeris and atomic clock errors, generates satellite upload of User navigation data, and maintains a record of satellite navigation processor contents and status. The MCS also has interfaces with the Satellite Control Facility (SCF) and Naval Surface Weapons Center (NSWC). The SCF provides a backup upload capability in case of ULS failure. The SCF also provides satellite telemetry and command information. NSWC generates a predicted ephemeris reference from MCS smoothed pseudo-range measurements for use by the MCS in the ephemeris estimation process. Figure 2 shows the interfaces between the satellites, User equipment, and the Control Segment.

Ephemeris and Clock Offset Computations

The satellite position and clock offset data employed by the User to compute his navigation solution is completely generated by the MCS. Figure 3 illustrates the operations of the MCS to perform this function. As seen from Equation (2), the MCS requires a time reference (referred to as GPS time). This time reference is established by the MCS designating one of the Monitor Station's atomic frequency standards as GPS time. GPS time is directly related to Coordinated Universal Time (UTC) by synchronizing an MS clock to the UTC calibrated "flying clock". However, GPS time does not contain the instantaneous 1 second changes required by UTC. A detailed discussion of the GPS time and estimation process described herein is given in a following chapter.

AVERAGING TIME (SEC)

Fig.13 NTS-2 cesium beam freqency standard performance

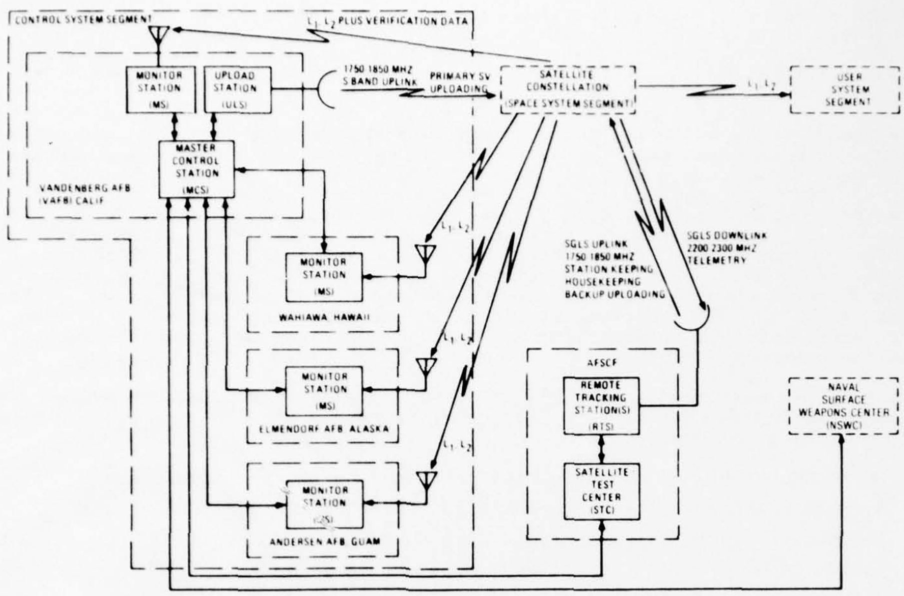
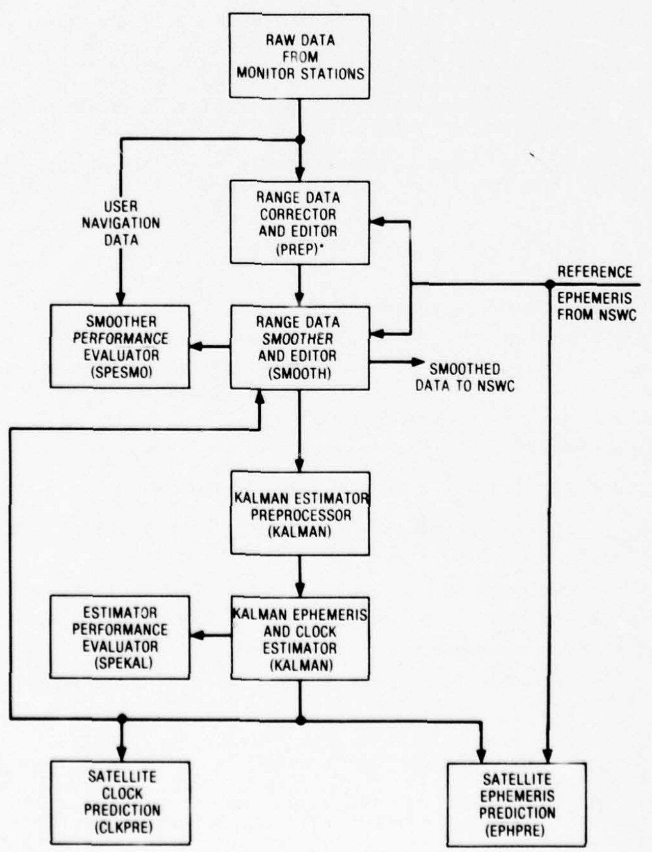


Figure 2. Global Positioning System, Phase I



*MCS SOFTWARE TASKS, SEE FIGURE 6

Figure 3. Ephemeris and Clock Processing

During a period of normal tracking (i.e., period not selected for satellite navigation processor loading) the MCS polls the Monitor Stations for raw data on nominal 5-minute intervals. The pseudo-range, delta pseudo-range, and meteorological data from the MSs are transmitted to the MCS over leased telecommunications channels. The ranging data is collected at nominal 6-second intervals by the MSs.

The pseudo-range and delta pseudo-range measurements are corrected and edited for wild points upon entering the MCS. The corrections made include time tag corrections to GPS time, antenna lever arm offsets, receiver calibration values, ionospheric delay, general relativity corrections to satellite clocks, and tropospheric delay corrections based on the MS meteorological data. The corrector also removes the reference ephemeris from the ranging measurements prior to smoothing (i.e., primary orbit kinematics are removed prior to smoothing). The smoother operates on either a 5, 15, or 30-minute interval of 6-second range data points. The editing of this data is based on three-sigma predicted measurement covariances from the Kalman estimator. Since the Kalman estimator is linearized about the NSWC reference trajectory, the reference must be added onto the smoothed ranges prior to transmission to NSWC. The reference ephemerides are not used again in the estimation process.

The smoother performance evaluator (SPESMO) computes parameters to evaluate data measurements during the normal data collection process. After each smoothing interval, the following data is computed:

- Differences between MS navigation solution and known MS positions
- Average measured ionospheric delay and ionospheric model differences
- Meteorological sensor measurement averages
- Residuals of MS measurements and smoothed pseudo-ranges
- Differences of smoothed and predicted pseudo-ranges

The smoothed ranging measurements (less reference ephemeris) are the primary input to the recursive Kalman estimating process. The Kalman processor is functionally separated into two parts: the pre-processor and the estimator. The pre-processor basically establishes the mathematical model configuration to be processed by the estimator. The primary function of the pre-processor is to establish the satellite partitions prior to estimation. The partitioning of satellite solutions results because of the tradeoff between testing the importance of correlated error sources and knowing that the number of states can become mathematically cumbersome. Therefore, the number of satellite solutions per partition is limited to any number less than or equal to four. The pre-processor also establishes initial conditions and reinitializes the Kalman estimator.

The Kalman estimator provides an optimum best estimate of the present satellite ephemeris and atomic clock error from GPS time. The individual states estimated for each portion are:

- Satellite position
- Satellite velocity
- Solar pressure constants (three components)
- Satellite clock bias, frequency offset, and drift rate
- Clock bias and frequency offset for three Monitor Stations (one MS clock is GPS time)
- Tropospheric residual biases for all MSs
- Polar wander (three states)

In the four satellite cases, a partition will contain up to 61 states.

The estimator performance evaluator (SPEKAL) provides visibility into the quality of the Kalman estimation. This is accomplished by providing:

- Residuals between reference ephemeris and Kalman estimates of satellite position
- Kalman states and their estimation error standard deviations
- Residuals between smoothed measurements and Kalman estimates
- Residuals between present satellite clock offset estimates and predicted values

The data processing described is typically performed during each Kalman estimate interval. The expected accuracy of these predictions is 3.66 meters (one sigma) for line-of-sight measurement of satellite ephemeris after 24 hours, and 9 ns of satellite clock error from GPS time after 2.5 hours. The elapsed time period for this accuracy is measured from the time the satellite navigation processor is uploaded. To complete the User requirements resulting from Equation (2), the MCS must use these predictions to generate a compatible data message that can be provided to the User via the satellite in a timely fashion.

Uploading Satellites For User Navigation

During the Phase I User testing, the satellites are uploaded at least once per day. Multiple uploads of User navigation data can also be performed to support special testing requirements. This process will be described in a subsequent paragraph and is only summarized here. During the upload process of six satellites, the CS (and in particular the MCS) is operating at nearly full capability. In addition to collecting ranging data from Monitor Stations and performing Kalman estimates, the MCS must generate satellite uploads, provide them to the ULS for transmission, and verify the satellite upload process by collecting telemetry (TLM) verification data from the Monitor Stations. This TLM verification data is contained in the User navigation data modulated onto the spread-spectrum signal by the satellite. The verification process is depicted in Figure 4.

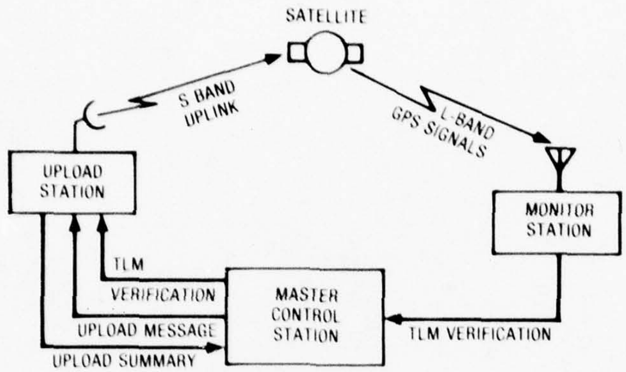


Figure 4. - GPS Satellite Upload Verification Process

The nominal satellite upload scenario consists of uploading the satellite just prior to its entering the testing area at Yuma Proving Ground (YPG). Depending upon the accuracy of satellite clock data desired for testing, the new User navigation data can again be provided to the satellite during the test period. Because of the importance and time criticality of the uploading process to User testing, the priority of satellite uploading in the Control Segment is much higher than any other activity. Uploading also requires resources such as devoted MS receiver channels for verification and MCS computer assets for generating upload messages and verifying the satellite upload process. Therefore, during the upload period, normal range data collection may be delayed and total CS resources must be allocated accordingly.

There are two navigation message uploads used to support User testing. They are referred to as the 6-hour and 26-hour upload. The 6-hour upload consists of satellite ephemeris and clock error parameters for only 6 hours. The 26-hour upload contains sufficient data for 26 hours of satellite orbit; and consists of satellite ephemeris, satellite clock, all satellite almanacs, and a special message. Either upload can contain satellite navigation processor diagnostic requests. The 6-hour upload will be used to initialize or refresh User navigation data for testing. The 26-hour upload supports once-a-day uploading, and final uploading before loss of satellite visibility to the ULS.

The upload process is completely controlled by the MCS. As shown in Figure 5, the MCS generates the upload message, maintains an image of the satellite navigation processor, monitors the uploading process, and verifies the satellite's transmission of the User navigation message. The upload generation is initiated and completely controlled by the satellite processor manager (RAMMAN). RAMMAN initiates the generation of predicted ephemeris and clock parameters for the upload. Almanac, special messages, and diagnostics are added if required. RAMMAN then selects the portion of processor memory to be loaded, and generates the User navigation message format tables. These tables control the duration and starting time of individual data messages. The upload generator (GENUP) formats the upload message into a satellite-compatible data form. The transmission controller (SENDUP) then sends the data to either the ULS or SCF for normal or backup uploading of the satellite, respectively.

The upload data is checked for transmission errors during the ULS upload process and upon transmission via satellite to the User. The upload verification is accomplished in the MCS by collecting TLM words from at least one MS receiver channel (TLMHOW). These words are presented to the MCS operator, and relayed to the ULS for the formal verification of the satellite uploading. The navigation message radiated by the satellite to the User is also systematically checked by the MCS every time it is changed by the satellite (NAVCHK). This data is collected by the Monitor Stations either automatically or upon command from the MCS, and relayed to the MCS for verification when requested.

The MCS must be able to support the satellite ephemeris and clock estimation, and satellite upload processes. Providing the navigation data to the User is of course of primary concern, as it is necessary for a navigation solution.

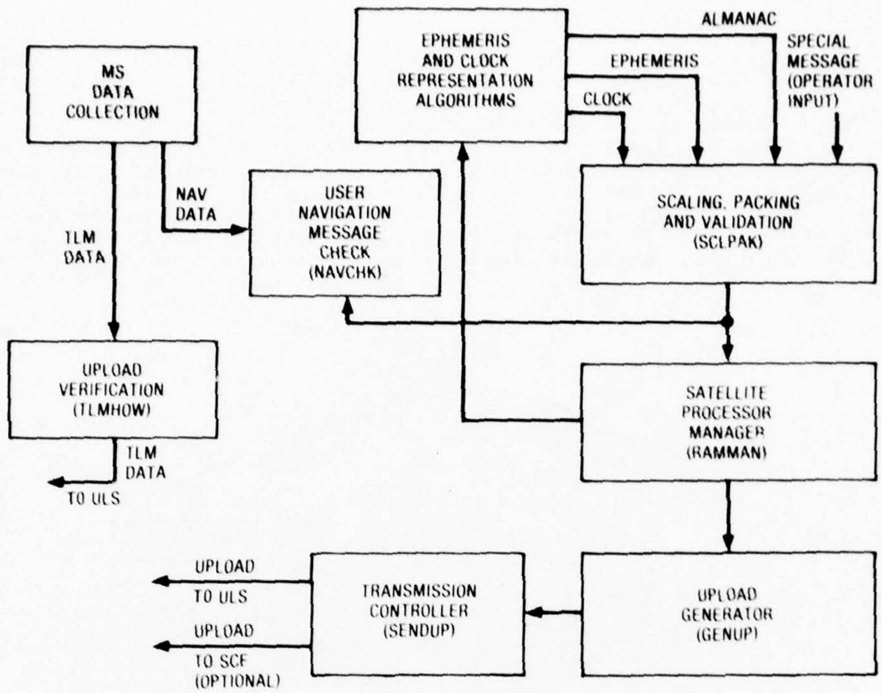


Figure 5. MCS Upload Message Processing

MCS SOFTWARE MECHANIZATION

The MCS software architecture is one of a collection of distinct software tasks isolated from one another by the data base. The system is run in a multi-tasking environment, with tasks vended to the operating system (OS) after determination that task precedence has been satisfied. Thus, as many tasks are running (up to a fixed limit) as can be running concurrently, considering the technical requirements of each task. To distinguish and control successive taskings, up to three byte-level parameters are passed to the task at commencement of execution. These parameters are often indexes to the data base, typically an MS or SV index. Each task is assigned a unique priority, and is allocated computer resources accordingly. Thus, tasks in the upload string (generally) outrank other tasks within the system, and command available resources during this time-critical phase of the operation. The operator can bump this priority (either for the prevailing execution or permanently) as the need arises.

The data base is the minimal-capability, disc-resident data base management system (DBMS) required to ensure data integrity and security. Data definition (i.e., the structure of the data base) was defined in conjunction with system synthesis and is fixed in the code. Data manipulations are restricted to simple GETs and PUTs into the data files for either data or header information (GETHDR/PUTHDR). With the exception of the three byte-level parameters, all input/output (I/O) passes through the data base. Data integrity is typically ensured through task precedence: The generating task shall complete before the using task is executed, and shall not be retasked before the using task has terminated. Where more than one concurrent task shares files, a simple queuing scheme augments task precedence. With the exception of shared tables, the DBMS code is local to the task. Figure 6 (although rather busy) illustrates this system-level data interdependence. Although most of the tasks are shown, only the critical files are represented on the figure. There are 70 distinct files in the MCS.

Security of the constantly updated data base is ensured by maintaining a "node" or kernel from which the operation may be restarted. This node is snapped upon any task's termination (providing its files are to be noded) by copying any portion (sector) of its data base which has been altered since its last execution. This alteration is maintained in a "sector vector" updated for selected files in the system by DBMS. Not every file must be noded, and the disc-resident data base is split between the noded and non-noded files.

Upon noding, the sector vector is reset awaiting the next task execution. Every noding is also output to the current system tables (principally the tasking stack) so that (upon recovery) any task either in execution or awaiting noding may be put back into execution upon recovery from the node. Thus the node may be described as an incrementally-updated data base, each increment consisting of those file segments altered by a given task following its successful completion. Actually there are two nodes for the recovery data base to eliminate: 1) problems occurring during the noding process; and 2) slow degradation ("poisoning") of the data base through non-fatal faults. Whereas the first

node is processed asynchronously as tasks complete, the second node is processed from the first at relatively long intervals (e.g., hourly) via a sector vector maintained between the first and second nodes. The two nodes are also located on separate removable disc packs, so that either node (providing it is viable) can be used for recovery even with one disc (of two) down. A portion of the data base is additionally contained on magnetic tapes, two of which are typically mounted and a third unit assignable. Data on tapes are not backed up on the nodes.

Operational command and control is exercised through a 9600-baud direct-view storage terminal (DVST) CRT. Primary operations feedback is obtained through a 1200-baud event printer, which posts eventful output asynchronously processed by each task in the system. An audio alert may accompany any event, and any event may be used to abort the task. Either condition is settable (resettable) by the operator at the DVST. Any file within the system can also be viewed (dumped) or the data content changed via the DVST. The planning function (PLAN in Figure 6), graphing of technical output (GRAF in Figure 6), and setting (resetting) system time, are typical functions of the operations interface. The task control function ensures that the operations control task (OPSCON) is always active, and reinitiates it immediately following an abort. OPSCON is typically in a "wait" state awaiting operator input. Maximum visibility of the operation of the system is afforded the operator by enabling viewing and alteration of the system tables.

Figure 7 illustrates the host computer hardware, a Xerox 550 with 128k 32-bit words of core (645 ns cycle time), two 11.8M bytes (8 bit) disc drives (each with 5.7M bytes on removable packs), three 800/1600 bpi magnetic tape units, two Tektronix 4014-1, 9600-baud DVSTs with a single (shared) hardcopy unit, three 300-baud alphanumeric data entry CRTs (only one of which is used for operational support), a 110-baud computer control TTY keyboard printer, a 300-lpm line printer, and a 200-cpm card reader. The line printer is only incidently used in operational support, the card reader not at all.

CONTROL SEGMENT SOFTWARE LEGACY

The software is interfaced with the host computer hardware/OS (currently the Xerox 550 and Computer Program-Real time, CP-R), through the system "core". The core contains all of the host computer-dependent code, and thus isolates the remaining code from host changes. This provides exceptional legacy for the bulk of the code as only the core (some 15% of the 60,000* lines of code) is host dependent. Specifically, assembly language instructions as well as Non-ANSI FORTRAN code is restricted to reside only in the core. As an illustration, code checkout of most of the code (except the core) was conducted on a CYBER 70 Model 172. Except for differences attributed to word length, results were identical.

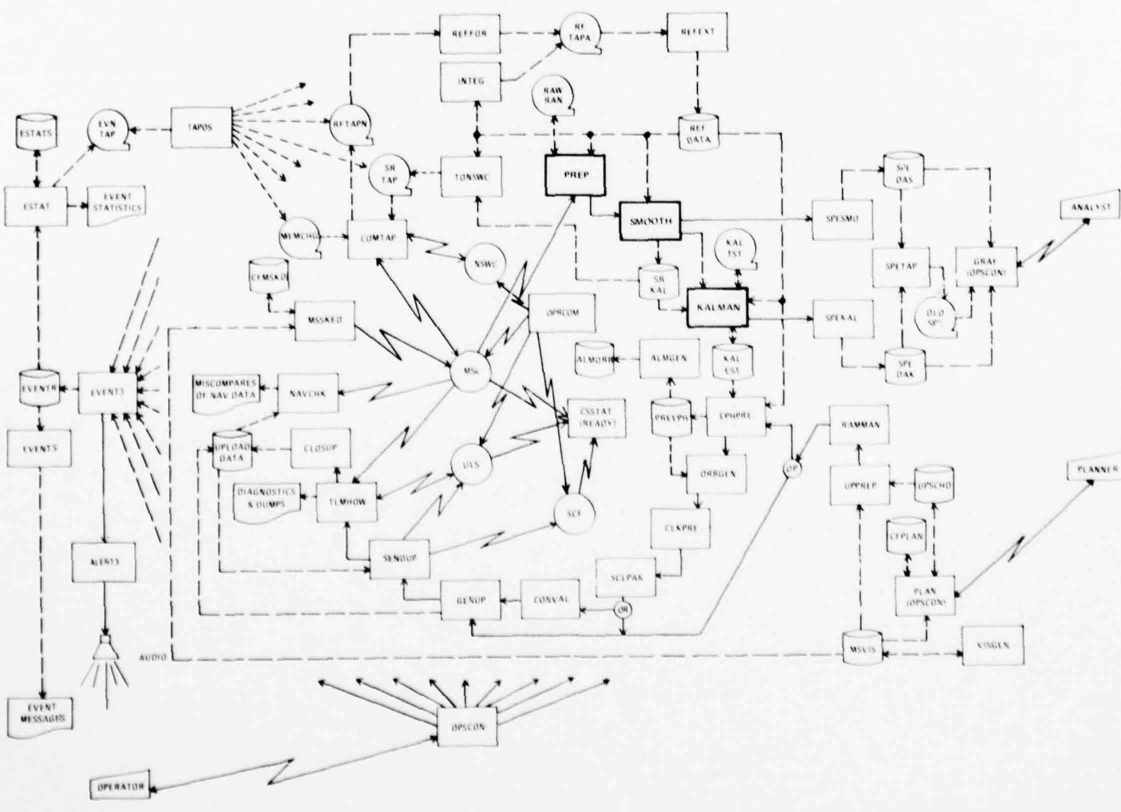


Figure 6. System-Level Data Interdependence Diagram

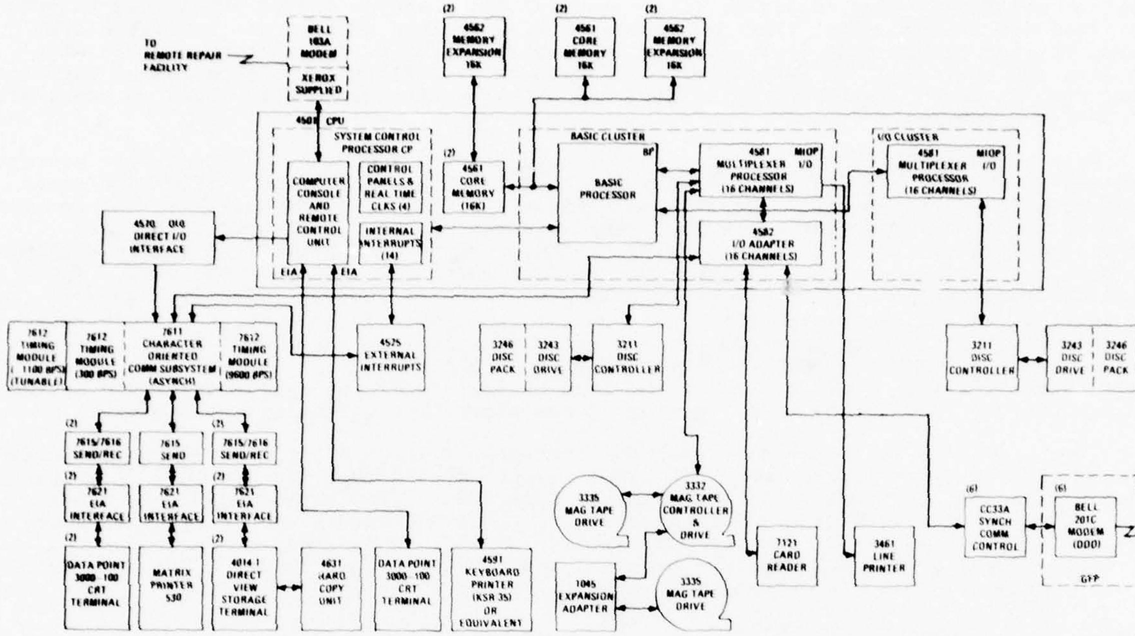


Figure 7. Master Control Station (MCS) Functional Block Diagram

The data base is structured such that it may be easily expanded to encompass additional remote stations (e.g., MSs) and SVs. Internal code buffers are commented with the algorithm defining the dimensions; a re-compilation following re-dimensioning would be all that is required.

The basic architecture additionally provides for the progressive offloading of operator duties as the level of support activities increases. The event outputs (appearing on the event printer at the operator's position) are designed to automatically invoke a software segment should corrective action be required. To the extent that the event parameters (groups of four words) together with their catalog number are definitive, additional decision processing can be encoded and a code segment added to the event processor. Thus, an operator decision and reaction which has become routine and tiresome can be relegated to the system with a relatively simple code modification. Further, routine operator procedures may be cataloged and stored, then later invoked in a hands-off fashion to ease the operator load and provide fast, accurate execution of relatively complex procedures.

The Control Segment software was developed with Phase III in mind, both operationally and to minimize host-hardware/operating-system change impact.

---

*This count includes only executable code (i.e., COMMON DIMENSION, comment cards, etc., are omitted from the count).

GPS MASTER CONTROL STATION OPERATIONS

H. C. Johnson
General Dynamics
Electronics Division
P. O. Box 81127
San Diego, CA 92138
U.S.A.

SUMMARY

The primary mission of the Control Segment is to maintain current and accurate navigational data on the RF (L-band) downlink of the satellites. To do this, the Control Segment tracks all of the satellites, takes range measurements, estimates the satellite's ephemerides, and periodically uploads current data and downlink-formatting instructions into the satellite's on-board computer. The satellite then downlinks the navigation data on the L-band RF signal according to the prescribed format. The navigation data is interpreted by the user system to determine the location of the satellite. This, together with the user's range measurements from four satellites, determines the user's position and time.

The Control Segment is designed to operate automatically in conjunction with an upload schedule that is manually entered. This suggests very straightforward and simplistic operating procedures. However, the functioning of the entire system in the "real-world" environment (basically, the environment for which it is designed) includes the operation of the hardware, software (or firmware), and system operators (controllers), and the interactions between each. The anomalous conditions that can arise from a combination of failures in any or all three categories suggest that complex algorithms are required to sustain operations and troubleshoot the system. This requirement is met by trained personnel with experience in operations and in the system (mainly software) design.

SATELLITE TRACKING

The satellite orbits are maintained very nearly circular with a period of one-half a sidereal day (altitude $\sim$ 20,000 km). Since the ground tracks repeat each sidereal day, so does the operational scheduling. The satellite visibility generator (VISGEN) is initiated (nominally) once each sidereal day. This visibility data, together with the satellite ephemerides, are used by the Monitor Station scheduler (MSSKED) to generate tracking polynomials. Tracking polynomials, together with tracking sequence data from the MCS data base, are transmitted to all Monitor Stations. The tracking parameters consist of a ninth-order-range polynomial and satellite rise and set times. The polynomial is used by the Monitor Station Computer Program (MSCP) to furnish the receiver with initial estimates of range and range-rate (doppler) to aid it in acquiring the satellite. The rise and set times tell the MSCP when a given satellite is visible. The tracking sequences basically assign any of the 24 satellites to each of the four receiver tracking channels. Normally, a satellite is assigned to a channel for a minimum of 2-1/2 minutes before another satellite will be assigned to that channel. If four or fewer satellites are visible, the satellites will remain assigned to the same channels until additional satellites come into view. Two different methods of sequencing are used; Monitor Station (MS) algorithm, or a Master Control Station (MCS)-generated sequence. The MS algorithm assigns the visible satellites according to an algorithm internal to the MSCP and in accordance with parameters furnished by the MS scheduler (MSSKED). These parameters include a start time that is normally used to synchronize tracking between the different Monitor Stations, a minimum time slot (normally 2-1/2 minutes, as described above), a channel de-allocation matrix, and a receiver polling interval. The polling interval is the time interval over which the receiver smooths the pseudorange measurements and integrates the doppler measurements. This is normally set at 6 seconds, i.e., the receiver will output one range measurement and one delta-range measurement every six seconds. The MCS-generated sequence assigns satellites to channels in a repetitive sequence. This sequence can consist of up to 16 different satellite/channel assignments that, upon completion, will repeat itself. This sequence also includes start and stop times such that one sequence can replace another, or, upon termination of the last valid MCS-generated sequence, the MS can revert to the MS algorithm. Once run, MSSKED automatically runs at least twice per sidereal day to generate new tracking parameters for each satellite/MS combination, as required.

The pseudorange tracking data collected by the MS is packaged by the MSCP in (nominally) 5 minute packages as designated by the MCS. Each package nominally contains as many as 50 measurements from one satellite (5 minutes at 6 seconds per measurement), and includes meteorological data for use in calculating tropospheric corrections.

The tracking data is retrieved from the monitor stations by the MCS tracking data preprocessor task (PREP). PREP is automatically tasked when the system is started up, and subsequently retasks itself to automatically collect data from all Monitor Stations. Each poll of a Monitor Station returns one package (nominally five minutes) of tracking data. When three such packages have been collected and preprocessed, it is smoothed (by SMOOTH) over the 15 minute interval and processed by the Kalman estimator (KALMAN). Thus,

the clock and ephemeris estimates are updated every 15 minutes. PREP also collects satellite (control-moment gyro) momentum dump data transmitted by the satellite and collected by the Monitor Station. This data is used by KALMAN to increase its uncertainty in the appropriate ephemeris states.

The PREP-SMOOTH-KALMAN string runs automatically, with no operator intervention. However, operator-initiated partial reinitialization of Monitor Station clock states, usually in conjunction with KALMAN backup, is sometimes employed when a Monitor Station clock reference has been lost due to power outage, equipment malfunction, etc. This procedure backs up the KALMAN to the estimates made in the (15 minute) interval just prior to the first valid data obtained from the Monitor Station in question (presuming that the problem has been corrected) and increases the uncertainty in the monitor station clock offset state for the next interval. This results in a new initial estimate of that state and the Kalman estimator (KALMAN) then proceeds in the normal manner.

UPLOAD

The Control Segment performs three basic upload types: Initialization; Control; and Navigation data. The second two may be combined. The initialization upload is done only once for a new satellite processor. The upload data block consists of the parity tables used by the satellite to put Hamming parity on the downlink telemetry (TLM) and hand-over-word (HOW) data words, and standard navigation data. All uploads will begin with an address block containing the appropriate satellite address and will end with an End-of-Message block that signals the satellite processor that the upload is completed.

A control upload contains time initiated events known as Z-counter events. These can consist of any or all of the available types of Z-counter events as described elsewhere in this document. A control upload consisting of a Z-counter adjust is always made following the initialization upload of a new satellite processor. This upload causes the satellite Z-counter to synchronize to GPS time. This upload will also be used any time the satellite clock diverges by more than 1 millisecond from GPS time.

A normal upload will consist of Z-counter events, Frame Formatter Table (FFT) cutover, modifications to the FFT, and optional navigation data processor diagnostics commands, two Frame Formatter Tables, and data. The data will consist of clock, ephemeris, and almanac data, and text messages. The clock and ephemeris data will normally be 6 or 26 1-hour "pages" predicted from the latest Kalman estimates of the GPS states. The almanac will contain 24 satellites' ephemeris and clock estimates with a dummy (25th satellite) filler of random numbers. The almanac data will normally be generated from the predicted ephemeris generated for that purpose during the day, but not necessarily using the latest Kalman estimates. This is because the almanac accuracy requirements are not nearly so stringent as the clock and ephemeris data. The text message may consist of up to five messages of 23 ANSCII characters each. These are downlinked one at a time by a single-page, repetitive FFT entry. The FFT is changed by SV processor memory modification Z-events at the time that the messages are to change. Since only five such events are included in each upload, the text messages may be "paged-through" only once, although the interval between pages is completely arbitrary. The first Z-event in a navigation data upload will consist of the first FFT cutover. Two FFTs are always used with the appropriate FFT cutover events. The second is (nominally) 15 minutes later than the first to allow for the contingency of uploading late and missing the first one. Since the second FFT consists of the first one propagated to the later time, the second cutover following a successful first is transparent to the user.

The navigation data processor diagnostics include processor logic tests, destructive memory (read/write) tests, and memory dumps. These are performed on selected areas of the satellite processor Random Access Memory (RAM) to diagnose possible problems, e.g., downlink parity errors that are detected repetitively, or at more than one Monitor Station.

Upload operations actually start at anytime with the interactive software task, PLAN. This task allows an operator to interactively select a candidate upload set consisting of an ordered set of satellites, times, upload types, and desired Z-events and text messages. PLAN then tests this candidate for conflicts and advises the operator accordingly. When an upload schedule has been found acceptable, the system task UPPREP is delay-tasked to run about fifteen minutes before the upload string is to run. UPPREP will check the schedule over for late-developing conflicts (e.g., a Monitor Station down in the system configuration file, CONFIG). If the schedule is correct, UPPREP then initiates the upload string for the upload set. The upload string then runs for each satellite in the set. The task SENDUP sends the data to the Upload Station (ULS) and tasks TLMHOW. TLMHOW then tasks MSSKED to assign the satellite to an MS receiver channel.

Upload verification is done by the task TLMHOW. This task retrieves the TLM and HOW data words received by the Monitor Station, sends them to the ULS for block verification, and polls the ULS for status. Upon successful completion of the upload, TLMHOW retrieves the upload summary from the ULS, builds the SUMARY file with it, and tasks CLOSUP. CLOSUP updates the MCS data-base copy of the satellite RAM, toggles the current/working copy pointers, and tasks MSSKED to take care of satellite navigation data downlink Monitor Station requirements (e.g., almanac, special message, diagnostic, and dump collection). The system task NAVCHK retrieves the almanac, text message, clock,

and ephemeris data, compares it to that uploaded, and notifies the operator of discrepancies. The task CSSTAT periodically (nominally every 5 minutes) checks the Monitor Station's status. The collection of clock and ephemeris page changes is, thus, detected by CSSTAT which tasks NAVCHK. MSSKED schedules NAVCHK for almanac and text message collection.

SYSTEM PERFORMANCE EVALUATION

The MCS computer system processes a large amount of tracking data and compresses it into an upload message. This whole process is monitored automatically by the System Performance Evaluator computer program (SPE). This program consists of several modules that execute after various steps of the data processing are completed. The SPE program provides the system operator with events statusing the quality of the processed data, as well as with major graphically displayed data in soft and/or hard copy.

A post-SMOOTH function module (SPESMO) generates parameters to monitor the performance of the smoother and related functions. It generates geometric dilution of precision (GDOP) values, computes differences between an MS location generated from a navigation solution utilizing received SV data and the known surveyed MS location, generates differences between the measured ionospheric correction and the correction computed from a single frequency user's ionospheric model, and computes the difference between the measured smoothed range and the range predicted by information from the SV NAV message. These values along with meteorological sensor values, the standard deviation of the range-smoothing process, and the actual measured ionospheric corrections are time tagged and stored on a data base file for use in trend plots and point analysis.

A post-KALMAN function module (SPEKAL) generates parameters to monitor the performance of the Kalman process and related functions. It generates differences between current estimates of SV clock parameters and those parameters predicted from the SV NAV message and those parameters estimated earlier, and it computes differences between current estimates of SV ephemeris parameters and those parameters estimated earlier. These parameters, along with the Kalman states and Kalman measurement residuals and their standard deviations, are time tagged and stored on a data base file for use in trend plots and point analysis.

A function module (SPETAP), upon operator request, merges the SPE data files onto magnetic tape. This tape is saved as a permanent record of all the data generated by SPESMO and SPEKAL and used for trend analysis.

The operating system data dump task formats all parameters stored by the SPE function modules and lists them in readable form on a line printer or direct view storage tube. The operating system data graphics task (GRAF), upon request, plots any or all parameters on the SPE data files on a DVST for trend analysis.

Conclusion

In conclusion, the operation function is to initiate and monitor the Control Segment equipment and software, and take those actions necessary to ensure that the data supplied to the user is of the highest quality.

MONITOR STATIONS

R. L. Harrington
General Dynamics
Electronics Division
P. O. Box 81127
San Diego, CA 92138
U.S.A.

W. A. Fabrizio
General Dynamics
Electronics Division
P. O. Box 81127
San Diego, CA 92138
U.S.A.

E. Salinas
General Dynamics
Electronics Division
P. O. Box 81127
San Diego, CA 92138
U.S.A.

SUMMARY

The Monitor Stations (MS) are fixed tracking stations in the GPS Control Segment which collect space vehicle (SV) tracking data, downlink navigation message data, and station health and status, under control of the Master Control Station (MCS). The MCS uses the tracking data to estimate and predict precise ephemerides of the GPS satellites. The clock and ephemeris predictions are used to generate the orbital element and clock parameters contained in the navigation messages for each satellite, which are uploaded daily. The functional relationship of the Monitor Stations to other parts of the system is depicted in Figure 1. The GPS Phase I (Concept Validation) Monitor Stations are located at Vandenburg Air Force Base (near the MCS), Hawaii, Guam, and Alaska.

INTRODUCTION

The Monitor Stations are housed in 7-foot by 12-foot military-type shelters* (Figure 2). The MSs are normally controlled entirely by the MCS, to which they are connected by dedicated communications lines, and are unattended.

The key component of the MS is the Set X GPS receiver, a high-performance, four-channel receiver developed by Magnavox. The receiver is identical to the User Set X except it does not have a Control/Display Unit, and the data processor is supplied with a computer program developed specifically for the Monitor Station.

One of the basic functions of the Monitor Stations is timekeeping for the system. The MSs are equipped with cesium beam frequency standards for this purpose. GPS time is established by the Master Control Station. Normally, the Monitor Station located at Vandenburg Air Force Base is the master Monitor Station; however, the MCS can designate any of the four MSs.

* Except the Alaska MS, which is in a pre-existing building.

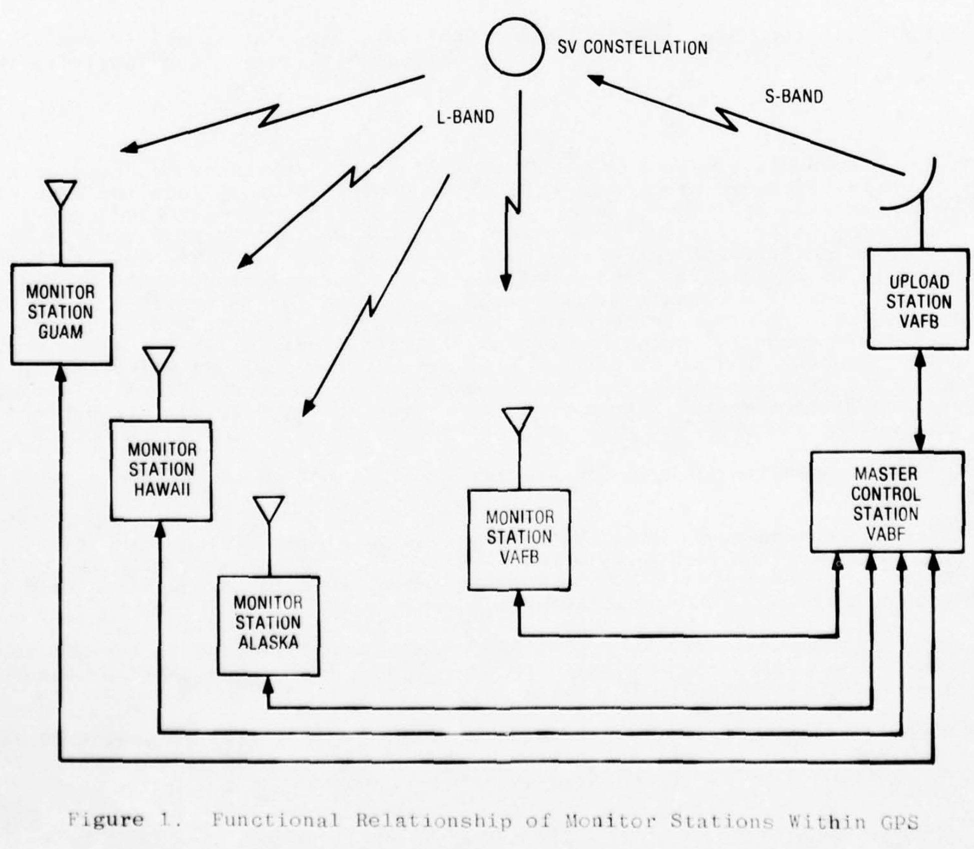


Figure 1. Functional Relationship of Monitor Stations Within GPS

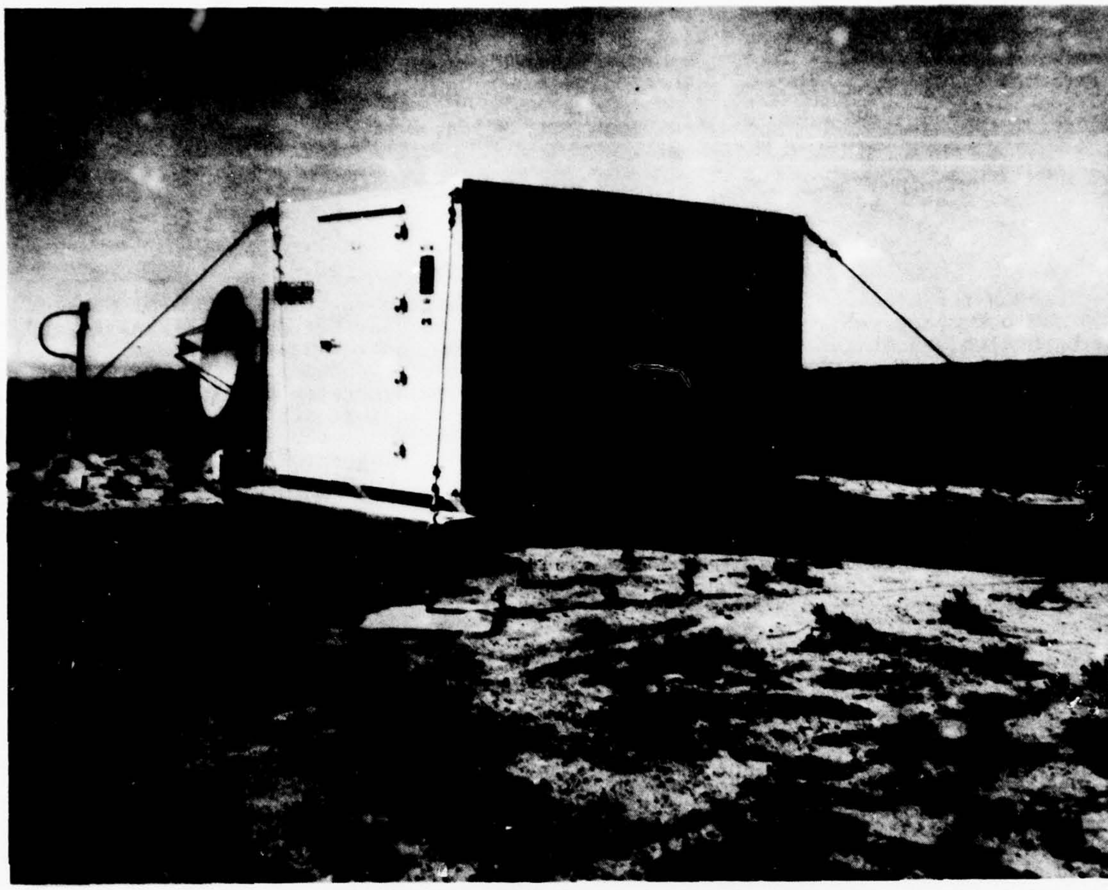


Figure 2. Vandenburg AFB Monitor Station

This paper describes the Phase I Monitor Stations; the operational (Phase III) MSs will differ in several respects, but the functions will remain substantially the same.

Block Diagram

The major components of the Monitor Station (Table 1) are interconnected as shown in Figure 3. Satellites are tracked on one of two L-band antennas; one for high elevations (30° to zenith) and the other for low elevations (5 degrees to 30 degrees), each with its own preamp. The receiver is controlled by the Hewlett-Packard (HP) 21 MX computer, which is also the recipient of all data from the receiver. The receiver tracking data is stored in an 8k-word extended memory buffer and transferred to the MCS on command. At the maximum rate of data accumulation, the 8k buffer is sufficient for about 20 minutes of data. If it is not collected by the MCS in this time, the data is written onto the read/write cassette, which has a capacity sufficient for at least two hours. The read/write cassette is also capable of storing computer programs which may be sent from the MCS. The programs can be loaded and executed on command. Receiver navigation message data and station status information are stored in fixed format messages in the HP 21 MX main memory and sent on request to the MCS.

The read-only cassette is used for program load. A reload can be commanded from the MCS.

The external environmental data system provides local meteorological data (temperature, pressure, and humidity) which is stored in the computer with the receiver tracking data. It is used by the MCS to compute corrections for propagation delays in the troposphere in the pseudorange and delta range data.

Internal environmental data is collected, stored, and forwarded to the MCS for monitoring purposes. Appropriate alarms are provided to flag undesirable or hazardous conditions, such as excessive temperatures.

The teleprinter is primarily for maintenance. There is a limited amount of monitoring and control of the operational system available through the teleprinter. It is also used for setting time in the MS from an external source.

The cesium beam frequency standard drives the receiver clock, which provides the time base in the computer. The pseudorange measurements and all time tags are based on this clock.

Communications with the MCS are made over a dedicated half-duplex line with C2 conditioning. Bell 201C (2400 bps synchronous) modems are used.

TABLE 1. Monitor Station Components

Item	Quantity
HP 21 MX computer with 40k-word memory	1
Cesium beam frequency standard with standby power supply	1
Cassette recorders	2
Teleprinter	1
Environment data systems	2 (Internal and External)
Signal conditioner (for environmental data)	1
Receiver system	1
Dual-element antenna	1
Preamplifiers	2
Receiver	1
Receiver power supply	1
28 Vdc power supply	1

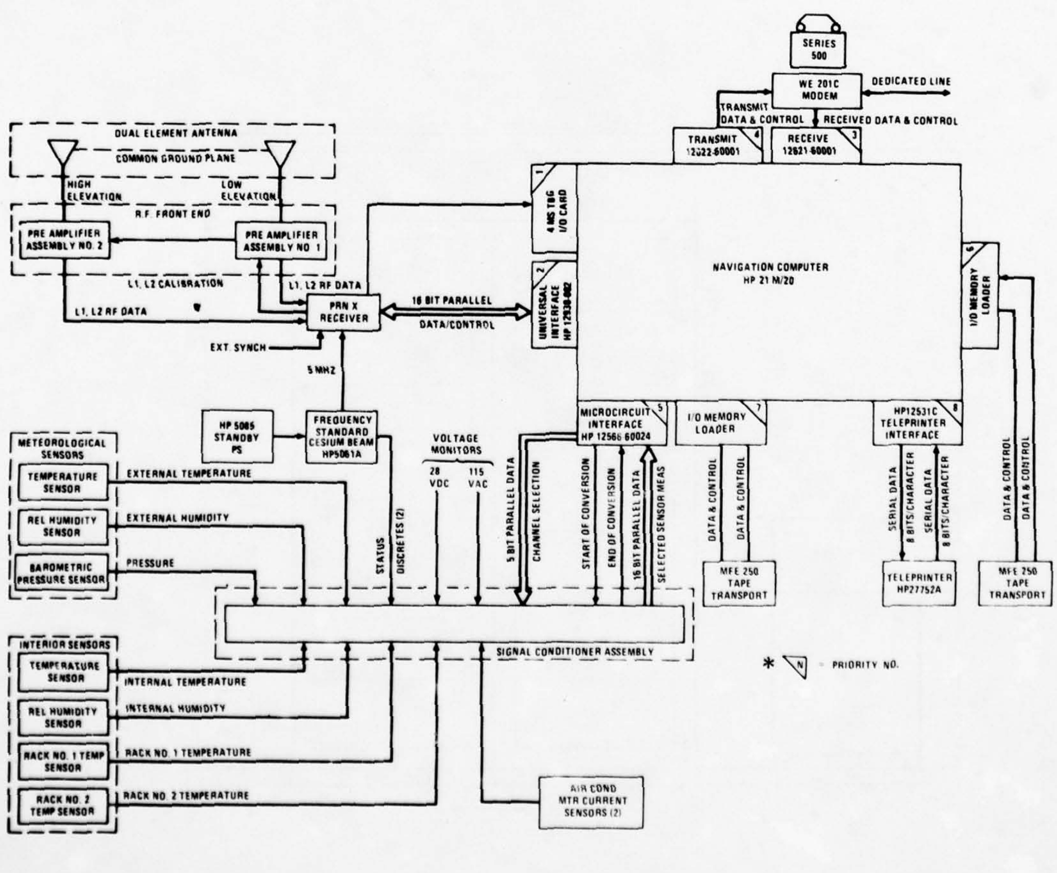


Figure 3. GPS Monitor Station Block Diagram

Equipment Layout

The electronic equipment is mounted in a double cabinet as shown in Figure 4. The antennas, a monopole for low elevation and a bent turnstyle for high elevation, are mounted concentrically under a small radome in the center of the roof of the shelter. A ground plane is provided. The shelter layout (Vandenburg, Hawaii, and Guam stations) is shown in Figure 5.

The Alaska MS is housed in a preexisting building. The equipment cabinets are as shown in Figure 4, except the antennas are mounted remotely on the roof of the building under a special 4-foot-diameter heated radome.

Monitor Station Satellite Measurements

The basic function of the Monitor Station is to collect pseudorange data on all GPS satellites. The measurements are performed by the 4-channel Set X User receiver, which is controlled by the Monitor Station Computer Program (MSCP). The receiver also measures delta range, which is an integrated doppler cycle count over a specified measurement interval, and an L1 - L2 pseudorange measurement, which is the difference in pseudorange between the L1 (1575.42 MHz) and L2 (1227.6 MHz) navigation signals emitted by the satellites.

The Set X receiver actually has five channels; four are used as carrier trackers and the fifth is time-shared for code tracking. The four carrier trackers are mechanized as Costas loops; they measure delta range and perform data demodulation. The code channel is a sequential, noncoherent, code-tracking loop shared among the (up to) four satellites being tracked. In addition, the code loop tracks L1 - L2 for each when so commanded. The code loop is rate-aided by the carrier loops. The receiver is capable of tracking four satellites simultaneously, each on either L1 or L2, whichever is specified.

The pseudorange measurement has a resolution of 1/64 of a P-code chip, or about 1.5 meters. The resolution and accuracy of the L1 - L2 measurement (used for ionospheric delay correction) are the same as for pseudorange. The delta range measurement resolution and accuracy are 1/64 of a carrier cycle (0.003 m) and 0.012 meter, respectively; these accuracies apply to P-code tracking, which is the normal mode for all MS measurements.

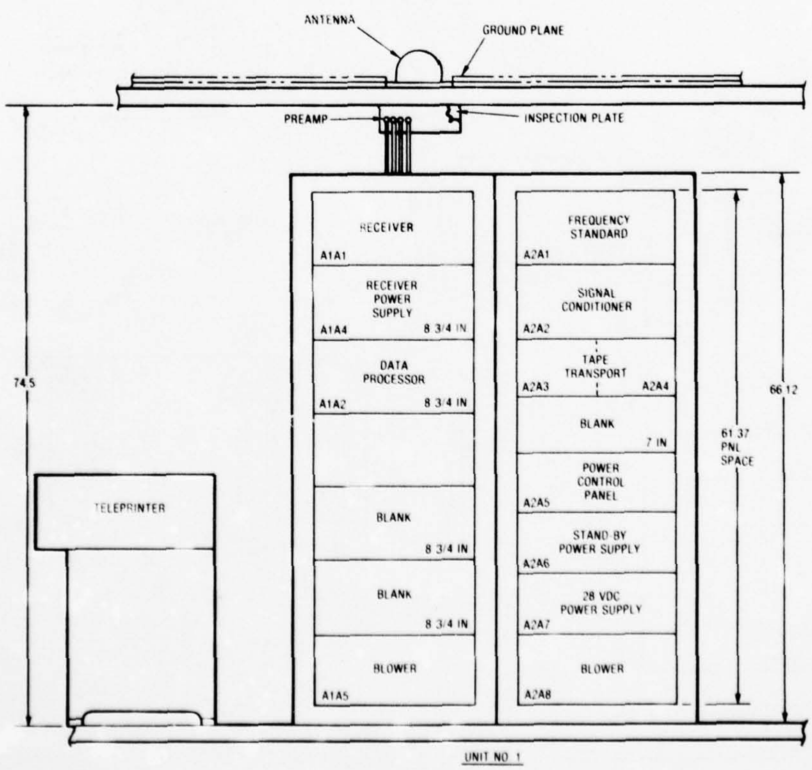


Figure 4. Equipment Rack Layout

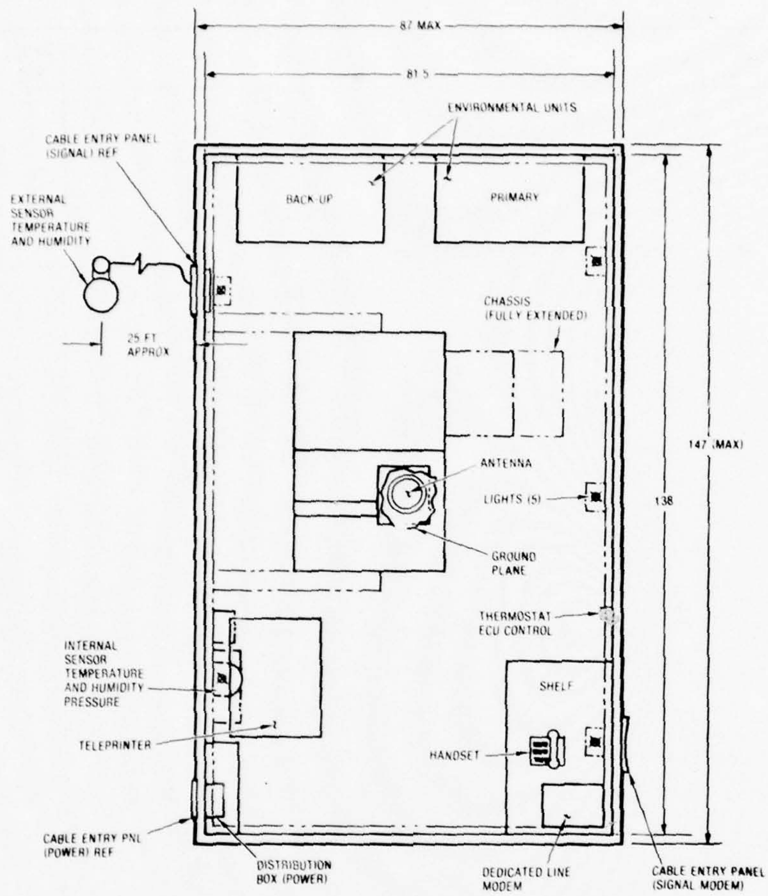


Figure 5. MS Shelter Layout

Meteorological Measurements

The Monitor Station measures the local meteorological environment including temperature, relative humidity, and barometric pressure. The surface environmental data is used by the Master Control Station to correct the pseudorange and delta range measurements for the tropospheric delay. The corrections are based on a model of the troposphere developed at JPL. The residual error between the actual tropospheric delay and the delay derived from the model is estimated by the Kalman filter. The residual error in the zenith direction, including the modeling error and sensor errors, is less than 10 cm over all possible meteorological conditions.

Control

The MS is message driven and under complete control of the MCS. The messages are listed in Table 2.

Message 01 provides station health and status including fault indications, alarms for various conditions, and presence of data of various kinds ready for transmittal to the MCS.

Assignment of the receiver channels is accomplished normally by a Message 27, which specifies the satellite and times over which it is to be tracked. Message 27 must be preceded by a Message 26, which contains SV visibility times, a ninth-order pseudorange polynomial used to compute a pseudorange and delta range estimate to aid acquisition, and antenna, code and carrier tracking assignments.

In the event that a Message 27 is not current, the MS tracks all visible SVs as defined by Message 26 data. If more than four SVs are visible, the MS rotates the channels among them using a predefined algorithm. The timing for data collection and channel time sharing is provided by Message 38.

Table 2. MCS to MS Request Message Responses

NUMERICAL IDENTIFIER	MESSAGE NAME	MS TO MCS RESPONSE
01	Station Health/Status Poll	01
02	Transmit SV Measurement and External Environmental Data	02
03**	Set Receiver Clock Using External Source	*
05	Transmit TLM and HOW Data	05
06	Transmit SV Navigation Message Parity Fault Data	06
07	Transmit SV Navigation Message Event Occurrence Buffer	07
08	Transmit SV Navigation Message Data Block I (change only)	08
09	Transmit SV Navigation Message Data Block II (change only)	09
10	Initiate SV Data Block III collection	*
11	Transmit SV Navigation Message Data Block III	11
12	Calibrate Receiver	*
13	Transmit Receiver Calibration Parameters	13
14	Clear Read/Write Cassette Directory	*
15	Transmit Read/Write Cassette Directory	15
16	Transmit Memory Dump	16
18	Transmit Read/Write Cassette Recorder Contents	18
19	Store Received Program on Read/Write Cassette	*
20	Load and Execute Cassette Program	*
21	Change Memory Command	*
22	Conduct Cassette Recorder Test	*
23	Reload and Initialize Command	*
24	Reinitialize Command	*
25	Set Receiver Clock from HOW Word	*
26	Tracking Polynomial Coefficients and Parameter Data	*
27	MCS Tracking Sequence Data	*
31	Transmit SV Navigation Message Data Block I (All SVs)	31
32	Transmit SV Navigation Message Data Block II (All SVs)	32
33	Initiate SV Navigation Message - Special Message Collection	*
34	Transmit SV Navigation Message - Special Message	34
35	SV Measurement and Environment Data Packaging Interval	*
36	Transmit SV and MS Multi-Message Data	36
37	Unassigned	
38	MS Tracking Sequence Scheduling Algorithm	*

*There is no message reply to these messages.
**Used only by teleprinter processor.

AD-A075 208 ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT--ETC F/G 17/7
PRINCIPLES AND OPERATIONAL ASPECTS OF PRECISION POSITION DETERM--ETC(U)
JUL 79 C T LEONDE
UNCLASSIFIED AGARD-AG-245 NL

3 OF 5
AD
A075208

The MS collects navigation message data from all of the satellites that it tracks. Data Blocks I (clock) and II (ephemeris) are collected and stored routinely. Data Block I and II data for all SVs are returned to the MCS via Messages 31 and 32, respectively. The MS continuously compares incoming data with the Data Block I and II in the MS data base. When a change is detected, the MS collects the data. The changes are reported via Message 01 and returned to the MCS via Messages 8 and 9. This system provides an efficient means of verifying the navigation messages on the satellite signals.

Data Block III (almanac) collection for a specific SV is initiated on command (Message 10) and the data sent to the MCS by Message 11. The special message is handled in a similar manner with Messages 33 and 34.

The telemetry (TLM) and handover word (HOW) are collected and made available via Message 5 at all times. The TLM word is used to verify the upload on a block-by-block basis.

The MS also monitors the downlink navigation data for parity errors which are reported in Message 06 and SV-related events (momentum dump and synchronization loss) which are reported in Message 07.

Setting MS Time

Time is set in the Monitor Stations by one of two methods; by use of an external source, or from the satellites. An external source is used to synchronize GPS to UCT (Universal Coordinated Time), to which GPS time is maintained to within 100 microseconds (neglecting leap seconds).

Under normal procedures, only the clock in the Vandenburg MS is set to the external source. This is accomplished by the use of cesium beam standard. The time of a future pulse from the external clock is entered on the MS teletype (Message 03) and receiver time is frozen at that value; when the pulse occurs, the receiver clock starts and is thereby synchronized.

The other Monitor Stations are set from the SV HOW in a manner similar to that used by User sets. This method has an uncertainty of 20 milliseconds. This offset is estimated by the MCS; once determined, the offset is simply a bias which is removed routinely.

MONITOR STATION COMPUTER PROGRAM

The MSCP is a real-time computer program composed of eleven tasks which are scheduled and operate on a priority basis. The tasks are programmed in MELTRAN (a structured language which is translated to FORTRAN) and HP assembly language.

A functional block diagram of the Monitor Station Computer Program is given in Figure 6. A functional description of each task shown in the block diagram of Figure 6 is given below:

a. An Executive task provides real-time multitasking processing support for task management and time-base generator interrupt handling for the MSCP. All tasks with the exception of the teleprinter, MCS/MS communications, and cassette recorder execute on a predefined sequence. These tasks are dormant until the next execution time matures. At this time, the task becomes active and executes on a priority basis. The remaining tasks are scheduled on task request only.

b. An Initialization task provides the start-up, restart, power fail recovery, and MSCP reload functions required to initiate or maintain MS processing. The initial start-up is a manual load of the MSCP from the read-only cassette recorder. Following the manual load procedure, the MSCP can be reloaded and initialized or reinitialized upon request from the MCS over the communication link. The initialization task supports system recovery from a temporary power fail condition. An extended power failure (greater than 2 hours) requires a manual reload of the MCSP.

c. A Communication task provides the message processing for communications between the MCS and the MS. A block diagram of the communication interface is shown in Figure 7. A checksum scheme is used by the communication software to check the validity of the message transmission. If the message fails this check on receipt at either end, a retransmit reply is given requesting retransmission of erroneous blocks.

d. The Receiver Controller task provides the control interface between the MS and the receiver process controller (PC) using the four fixed-format data groups described in Table 3. The receiver control parameters are computed from the SV tracking parameters provided by the MCS over the communication link. The tracking sequences for the four-channel receivers is also specified by the MCS. The tracking sequence provides SV assignments, start and stop times, receiver polling intervals, and a time interval that a specific set of SV channel assignments will remain in effect. The receiver controller provides a receiver time set and MS time synchronization capability.

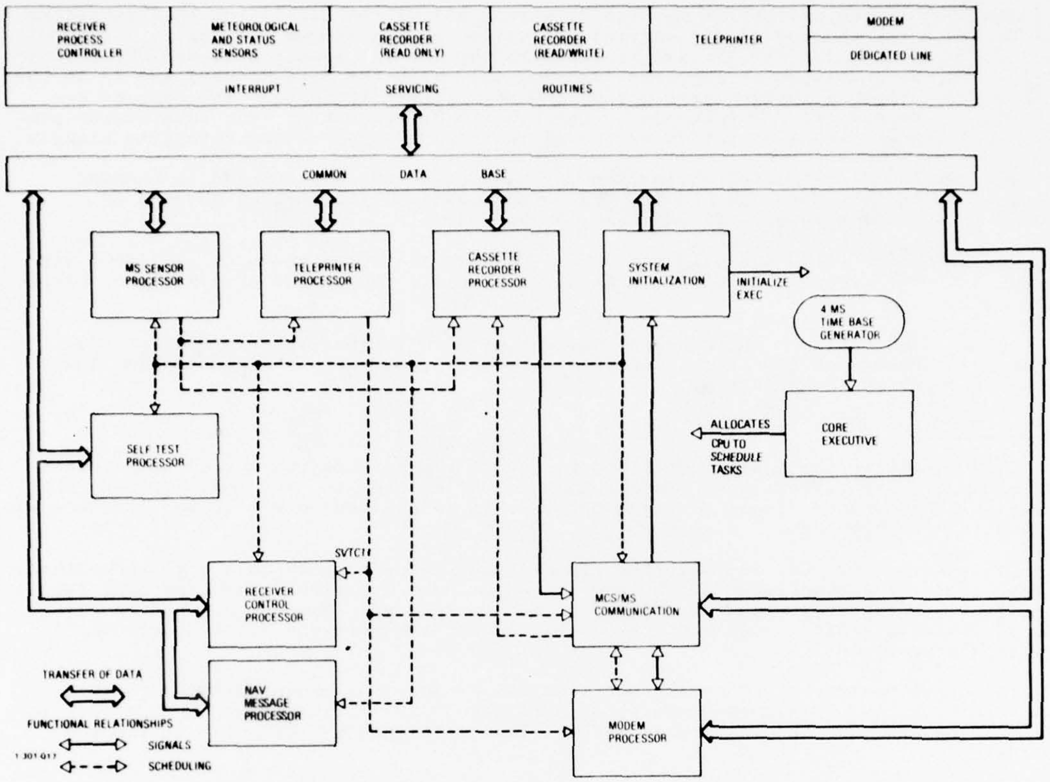


Figure 6. MSCP Functional Block Diagram

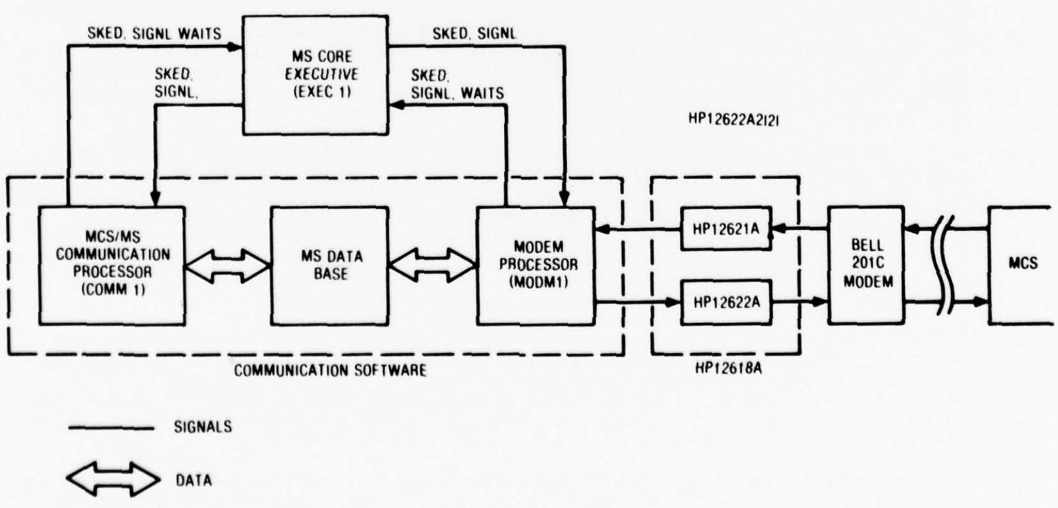


Figure 7. MS/MCS Communication Interface

Table 3. Receiver Data Groups

GROUP NUMBER	DP INPUTS AND OUTPUTS	DATA CONTENT DESCRIPTION
I	Input (27 16-bit words)	Pseudorange, delta range, channel health and quality, receiver quality, and user time-of-week.
II	Input (80 16-bit words)	Demodulated satellite navigation data and L1/L2 ionospheric delay measurement.
III	Input (128 16-bit words)	Receiver RAM dump.
IV	Output (23 16-bit words)	General receiver control, acquisition control, track control and override options, estimated pseudorange, estimated pseudorange rate.

e. A Navigation Message task processes SV navigation message data transferred to the MS from the PC. The SV navigation message clock and ephemeris data are continuously compared to data base values for each SV being tracked and collected on a change-only basis. SV navigation message almanac and special message data are collected for an SV only on request from the MCS. The navigation message task provides an SV health and status monitor function by detecting SV synchronization loss and SV roll momentum dump occurrence. This task also determines SV navigation message validity by detecting word synchronization errors, parity errors, and receiver demodulator failure. Telemetry and handover words are continuously collected to support MCS upload verification and SV processor diagnostics.

f. A Meteorological and Station Status Sensor task periodically samples environmental and MS station status data, and external environmental data and packages along with SV tracking measurement data for transfer to the MCS. The external environmental data is used by the MCS to correct the SV tracking measurement data for tropospheric effect.

g. A Self-Test task performs a continuous memory check and instruction test as a low-priority background function and reports errors to the MCS upon request.

h. A Cassette Recorder task provides for an operational reload from the read-only or read/write cassette, stores a minimum of one hour of SV tracking measurement data on a read/write cassette, and stores computer programs received over the communications link on a read/write cassette for load and execution at a later time.

i. A Teleprinter task provides an operator interface for communication with the MS computer program.

j. A Modem task provides the communications I/O interface between the MCS and MS.

<u>Conclusion</u>

The Monitor Stations described in this paper were built as part of the GPS Phase I system, the purpose of which was to validate the concept of GPS. This objective had been achieved in large degree by the end of 1977, with three of the four Phase I Monitor Stations in operation in conjunction with the Master Control Station and the first GPS satellite, NTS-2. Successful operation of the Monitor Stations had been demonstrated in all significant modes of operation, and the viability of the design was proven.

THE GPS UPLOAD STATION

W. H. Harbour
General Dynamics
Electronics Division
P. O. Box 1857
Vandenburg AFB, CA 93437
U.S.A.

M. Rochette
General Dynamics
Electronics Division
P. O. Box 1857
Vandenburg AFB, CA 93437
U.S.A.

SUMMARY

The GPS Upload Station (ULS) provides the Master Control Station (MCS) space uplink for the transmission of the navigation data payload to the GPS navigation satellites. The ULS receives upload data messages from the MCS as well as additional data necessary to point the ULS antenna properly at the space vehicle (SV) selected for upload. Data enters the ULS over a dedicated communication line and is received by the ULS computer. The computer restructures the data into SGLS-type signals and provides it to the ULS RF equipment where it is amplified and transmitted to the SV. The MCS provides the ULS with an indication of message receipt by the SV. Upon upload completion, the ULS prepares a summary of the upload results for the MCS.

INTRODUCTION

The GPS Upload Station is physically located in the same building as the MCS. Figure 1 is a photograph of the interior of the ULS showing some of the operational equipment. Figure 2 is a block diagram of the ULS. The reader may wish to refer to Figure 2 in the following description of the ULS hardware.

Upload Station Hardware Description

The Upload Station receives command and antenna pointing data from the Master Control Station and subsequently delivers uplink commands to GPS satellites through the use of computer, command transmitter, and antenna control subsystems.

Communications between the ULS and MCS is effected over a half-duplex dedicated communications line. The line operates synchronously at a 2400-baud rate through the use of Bell 201C modems. All upload data or requests for upload status originate in the MCS and are processed by the ULS computer.

Figure 1. View of the Operations Stations of the ULS

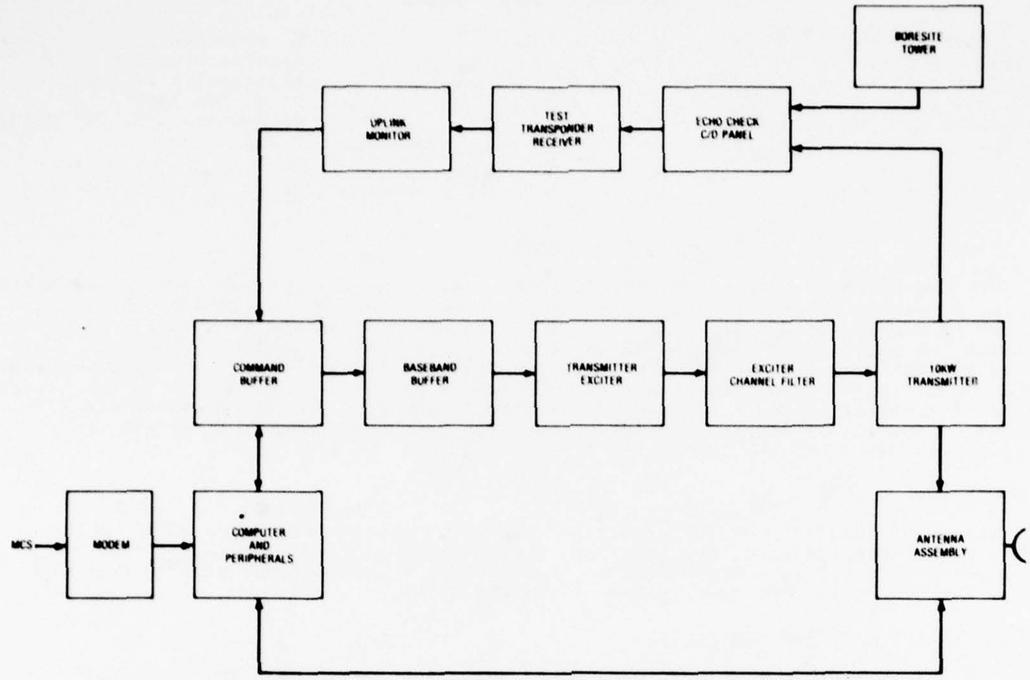


Figure 2. Upload Station Simplified Block Diagram

Computer Subsystem

The computer is a Hewlett-Packard (HP) 2108 computer which consists of the central processor, memory, and I/O. The computer features hardware power fail and automatic recovery logic, memory parity checking, memory protection, and direct memory access (DMA). Peripherals utilized by the ULS software are data communications modem, disc drive, system console, time-base generator, card reader, line printer, tape cassette, paper-tape reader, and an operator's console. The disc drive is an HP 12960A disc subsystem with a capacity of 2.4 million 16-bit words. The maximum DMA transfer rate is 616,666 words/second. The system console is an HP2762A terminal keyboard/printer with a 30 character/second printing speed and 128 character ASCII set. The time base generator is an HP12539C PCA which generates a pulse every 10 ms. The card reader is an HP12986A Optical Mask Reader Subsystem which reads 200 cards/minute. The line printer is an HP12987A dot-matrix printer which prints 200 lines/minute with a paper slew rate of 3.2 inches/second. The tape casette is an MFE model 250 digital cassette tape transport unit with a recording density of 400 bits/inch at a read/write rate of 10 inches/second. The paper tape reader is an HP12915A high-speed unit which reads at the rate of 500 characters/second. Lastly, the operator console is an HP2640 interactive display terminal.

The processing performed by this computer system is described in a later paragraph.

Command Buffer Interface

Upload information processed by the ULS computer is fed to the Command Buffer and Antenna Control Unit. Two interface cards transfer the digital command data from the computer to the Command Buffer. A third interface card counts and stores echo errors as they occur for real-time evaluation, and also delivers a command 1 kHz to the Command Buffer. Commands are supplied to the Command Buffer over parallel lines representing S, 0, and 1 characters. If no command is available, the Command Buffer supplies an S bit to the Baseband Assembly.

Commands from the Command Buffer are applied to the Baseband Assembly unit in the form of logic levels. The command logic gates an output from one of three frequency-shift-keyed (FSK) oscillators, generating a command subcarrier. The output of the FSK oscillators is 65, 76, or 95 kHz for an S, 0, or 1, respectively. The FSK output is applied to an amplitude modulator along with the 1 kHz command rate sync pulses that are first integrated to produce a triangular waveform. The amplitude-modulated FSK command signal is fed through a selectable attenuator where a modulation index of 0, .3, or 1 radian is selected before the composite baseband signal is sent to the modulator in the transmitter exciter.

The Transmitter Exciter provides a crystal-controlled, phase-modulated RF output of +10 dBm minimum at a frequency of 1783.74 MHz as a low-level input to the Exciter Channel Filter. The modulation index is adjustable to 1.5 radians and is set to 1.0 radian.

The Exciter Channel Filter suppresses all spurious and harmonic signals present at the RF input to a level of at least -60 dB and provides a controllable RF output level. The RF filtering is accomplished through the use of a Yttrium-Iron-Garnet (YIG) filter. The unit has a filter bypass mode.

The Exciter Channel Filter output drives an HP489A microwave amplifier, which is a broadband linear amplifier that provides RF signal amplification of at least 30 dB. The Traveling Wave Tube (TWT) utilizes periodic permanent magnet focusing. Maximum power output is approximately 2 watts (+33 dBm).

High-Powered Amplifier (HPA)

The Energy Systems Inc. model 11-160 Wideband Transmitter amplifies the low-level driver output to a power level of more than 10 kilowatts (+70 dBm), or amplification is accomplished through the use of a Traveling Wave Amplifier. The TWA is water-cooled, as is the solenoid (focusing magnet), waveguide circulator, turnable ferrite mode suppressor, and the dummy load. A 65-gallon heat discharger controls the water temperature. High voltage is provided by a 3-phase, 420 Hz motor generator. The 3-phase power is rectified and supplies 20,000 Vdc at 5 amps to the TWA. A crowbar circuit, actuated by high reflected power, false frequency, waveguide arcing or excess body current, prevents destruction of the TWA by shorting the high-voltage supply through a hydrogen-filled thyratron tube. The TWA output power is fed to the antenna through a high-power, low-pass waveguide filter which attenuates the spurious second and harmonic frequency signals which are generated by the TWA. The output power can also be fed to the RF dummy load.

The Test Transponder Receiver (TTR) input is fed from a directional coupler in the HPA. The TTR is a dual-conversion, superheterodyne receiver with an MDS of -80 dBm. The TTR supplies the demodulated uplink composite baseband to the Uplink Monitor.

The Uplink Monitor separates and demodulates the composite baseband signals and routes them to the Command Buffer as DC logic levels (S, 0, 1) where they are used for echo check purposes.

Antenna Control System

As the upload window approaches, a ULS computer-generated "Program Track" command gives the computer control of the Antenna Control Unit (ACU). Azimuth (AZ) and elevation (EL) slave angle data are transferred to the ACU and the antenna is pointed toward the satellite. Antenna position data is returned to the computer from the ACU. If angle errors of greater than two degrees occur, an AZ or EL alarm indicator is lighted on the ACU, and antenna position comparisons are continually printed out on the line printer until the antenna position error is corrected.

The antenna system has six principal operating modes: Standby, Manual Rate, Remote Manual (radome control), Program Track, Local Designate, and Auto Track. The primary operational mode is Program Track. In this mode, 12-bit scaled binary angle data is applied through the computer interface, stored in a binary comparator, applied to a digital to analog converter, then amplified and used as a field current which is applied to the control windings of each generator in the motor-generator sets. The outputs of the generators are then applied to a pedestal drive motor in each axis. Antenna position data from pedestal-mounted 1:1 coarse synchros is applied to synchro-to-digital converters in the ACU, the output of which is applied as the second input to the binary comparator and acts to null the error signal when the ordered antenna position is reached. The synchro-to-digital converter output also supplies the computer angle information for position error calculation and drives an ACU-mounted antenna position indicator.

The resolution of the antenna system is 0.089 degrees. Program Track mode tracking rates are 5.23 degrees/second in AZ and 0.5 degrees/second in EL.

Antenna

The antenna is comprised of a 14-foot parabolic dish, direct-drive dc motors for each axis, a slip-ring assembly to permit unlimited movement in AZ, and rotary joints for RF passage through the pedestal. Right-hand circular polarization is utilized for all operations. The antenna RF gain is approximately 31 dB at the operating frequency. The RF beamwidth is approximately 3.0 degrees in each axis. The antenna is housed in a rigid urethane foam radome 26.5 feet in diameter. It is equipped with normal and plunge and reverse optical collimation and Polaris sighting windows. Seventeen-bit optical encoders are used to align the antenna to a reference point.

Normal Software Operation

The purpose of the ULS is to upload navigation data to the SV. The message containing upload data is identified by the ULS as a prepass message. The prepass data sent from the MCS is composed of two parts. The first part is the acquisition data,

which contains the antenna pointing angles for the SV to be uploaded and the time the upload is to begin. The second part is the upload data that is to be transmitted to the SV. The prepass function separates the MCS data into an acquisition file and an upload file. Based on the time given in the acquisition data, the upload task is scheduled.

The upload task controls the transmission of the upload data to the SV. This task interfaces with the antenna pointing task to ensure proper antenna position and with the SV block verify task to determine the status of each block uploaded to an SV. The upload data is composed of blocks. There are three types of blocks to be uploaded. The first type is the address block. The address is SV specific and alerts the SV for data to be uploaded. The upload task continues to transmit the address block until the status set by the verify task has confirmed that the address block has been accepted by the SV. Once the address block has been verified, the second type of block, data blocks, are transmitted. If there is more than one data block, each is transmitted in sequence until statused as accepted by the verify task. When all data blocks have been accepted, the third and last type of block, the End-of-Message (EOM) block, is transmitted. Once the EOM block is accepted, the upload task is terminated. During upload task termination, a summary of the upload is created and statused as available for the MCS. It is availability of the upload summary data that starts the termination of upload monitoring at the MCS.

ULS Software Architecture and Tasks

The ULS software system is a real-time system supporting built-in test capability and a batch diagnostics capability. The major ULS software tasks are RTE-II, Control and Display (C&D), Prepass, Upload, Antenna Pointing, Verifier and Diagnostics. RTE-II task is an HP real-time executive.

The antenna control driver is activated by an RTE-II EXEC call to input the current antenna position, and an EXEC call to output a position command to the antenna control unit.

The data modem driver uses RTE-II privileged interrupt in responding to interrupts from the data modem. The driver is normally in the read state monitoring the communications line from the MCS. Only when the MCS requests data from the ULS does the driver transmit data to the MCS.

The upload driver initializes the command buffer which then generates interrupts on a 6 millisecond cycle until it is deactivated. After this driver is activated, it is controlled by the upload task. The upload task passes the driver the data to send and the number of bits of data in the buffer. When the data has been sent, the driver sends S-pulses until the upload task passes another address and bit count or sets the stop control variable. When the stop control variable is set, the upload driver deactivates the command buffer and dequeues.

A Control and Display (C&D) task is also provided. The C&D task responds to C&D operator requests, manages data, and control messages to and from the MCS and calls from other tasks. Operator responses permitted include modifying the SV upload time, deleting SV entries from the upload schedule, queueing, displaying, upload schedules, initiating type B diagnostics, aborting upload, displaying SV history, displaying block verify status, scheduling the next upload, verifying upload blocks, displaying SV upload summaries, setting/resetting block status display flags, and sorting the upload schedule in ascending order of upload times to ensure that the first entry is the next SV to be uploaded.

MCS message processing includes sending the upload schedule to the MCS, initiating a type B diagnostic, aborting an upload, sending an upload status summary to the MCS, sending ULS system time to the MCS, sending an upload summary to the MCS, rescheduling an upload operation for address and EOM blocks only, receiving prepass data, and identifying data line messages from the MCS and passing them to the appropriate software task.

Calls from other tasks include calls to display the upload schedule on the ULS line printer, schedule the next upload operation, manage data communications output to the MCS, display a summary of an SV upload, and sort the upload schedule in ascending order of upload times to ensure that the first entry is the next SV to be uploaded.

The prepass task is responsible for the creation of all ULS operations files in support of upload operations. This includes formatting an acquisition data disc file, an upload data disc file, upload schedule disc file, and command status table disc file. The prepass task also tabulates all new SVs for which files are created or updated, tabulates error messages as errors are encountered, and activates the C&D task to schedule the upload task when all data has been processed and all files created.

The upload task controls the transmission of the upload data to an SV. The upload task is called initially by RTE-II up to 30 seconds prior to the scheduled upload time. The upload task schedules itself through RTE-II to control the transmission of the upload message. The upload task processing proceeds as follows:

1. Initialize the upload and antenna pointing, and verify tasks prior to uploading an SV.

2. Initiate the transmission of address, upload data, and end-of-message (EOM) blocks.

3. Select the next block to be transmitted based on verification word received from the verify task.

4. Perform processing to initiate the next block after each block is transmitted.

5. Terminate the upload task following a successful EOM transmission or a successful abort. Alarm the MCS and ULS operator when an SV upload is not completed within the specified SV upload window, or the address block is rejected, or an operator abort is received.

The antenna pointing task positions the ULS antenna using azimuth and elevation polynomials provided by the MCS. Once these angles are verified to be within range, they are supplied to the antenna control unit. The actual antenna position is compared to the computed angles to verify correct antenna positioning within a specified antenna tolerance (currently set at 2 degrees). An error message is generated if the antenna is not positioned within tolerance.

The Verifier task is responsible for receiving upload block verification messages from the MCS and notifying the Upload task as to the status of a block upload operation. This is accomplished as follows:

1. Interpret the verification message from the MCS to determine if a block upload was accepted or rejected.

2. Set the block accept flag for blocks that have been accepted by the SV.

3. Increment the reject count for a block that has been rejected.

4. Compile the block verification status record and write it in the history file in disc each time a verification message is received.

5. Format the block verification record for display.

The Diagnostic task's role is to identify and isolate faults that affect the ULS system operation. There are three types of diagnostics: a) Type A are RTE-II alarms provided by the vendor software; b) Type B are operator checks of ULS hardware; and c) Type C are stand-alone diagnostics that are loaded and run independently of the ULS operating system.

The software just described has been programmed and installed on the Upload Station computer system, and the ULS has been delivered to the field. Currently, the ULS is uploading the navigation satellites on a daily basis in support of GPS Phase I field operations.

A TIME TRANSFER UNIT FOR GPS

Jackson T. Witherspoon, Leonard Schuchman
Stanford Telecommunications, Inc., Sunnyvale, CA & McLean, VA

SUMMARY

A Time Transfer Unit (TTU) for use with NAVSTAR satellites of the Global Positioning System (GPS) is described. This unit can provide worldwide time transfer with an accuracy of better than 100 nanoseconds relative to GPS system time. Techniques used for GPS signal processing and data reduction are summarized, and potential applications discussed.

1.0 INTRODUCTION

Satellites have been used for a number of years for transfer of time and frequency between primary standards and users of precise time (1). NAVSTAR satellites of the Global Positioning System (GPS) will soon provide world-wide access to precise timing signals which will permit a properly equipped user to calibrate a local time source relative to a primary standard to within fractions of a microsecond. The minimum user configuration consists of an omni-directional antenna and preamp, a Time Transfer Receiver, a teletype or equivalent alphanumeric printer, and a scientific hand calculator. Users with existing computers can eliminate the manual computation by connecting the receiver directly to a host processor.

For users without existing processing capability, an integrated receiver/processor or Time Transfer Unit (TTU) is described which is completely self-contained and fully automated. For users who do not have atomic frequency standards but who wish to improve the accuracy of their local time reference to better than 1 microsecond, the TTU can be operated in an automatic time correction mode in which the local time reference is periodically corrected.

This paper describes the operation and design of a prototype unit now under development.

2.0 OPERATION

The basic configuration of the TTU is shown in Figure 1. The unit accepts stable reference time (1 pps) and frequency (5 MHz) signals from the local time standard or unit to be calibrated, and the L1 signal at 1575 MHz from any GPS NAVSTAR satellite. Ephemeris data from the NAVSTAR satellite is detected and processed to determine satellite position and to estimate the time of arrival of the satellite signal epoch. The actual time of arrival of the signal epoch is recorded and compared with the expected value after correction for ionospheric, tropospheric, and relativistic errors, and the difference is displayed as the local time error. Sequential time error measurements are filtered to provide minimum variance time and frequency error data for display to the user or for logging with other user applications data. Error of local time is displayed as a nine digit plus sign decimal number in units of seconds with a range of $\pm$ six seconds and resolution of 10 nanoseconds. Frequency error is computed from the rate of change of the filtered time error and displayed as a six digit plus sign decimal number in parts per 10^9 with a range of $\pm$ one part 10^6 and a resolution of one part in 10^{12}. The output data is updated every six seconds; the normal smoothing interval is two minutes but can be adjusted to a particular user's requirements. The smoothing filter also generates an estimate of the variance of the output data.

Several optional modes of operation are available to match different user requirements. For the user who needs a corrected 1 pps time reference, a time synthesizer can be added to the basic configuration which generates a 1 pps which is advanced or retarded relative to the reference 1 pps by an amount equal to the computed time error.

2.1 Time Transfer Technique

The time transfer technique used is illustrated in Figure 2 which shows typical satellite, user and received time epochs relative to system (GPS) time. GPS satellites transmit continuous navigation signals with readily identified epochs every 6 seconds. The satellite transmission is determined by an atomic standard which will, in general, differ by some amount (Parameter A in Figure 2) from system time. An estimate of A is contained in the navigation data and is available to the TTU. The satellite epoch arrives at the TTU and is measured by reference to the local station time reference (Parameter C in Figure 2). The transmit time (sum of A + B + C in Figure 2) can be estimated from satellite ephemeris and ionospheric error available in the navigation data, and from known user location and equipment bias calibrations. The station time error is taken as the difference between the local time and predicted GPS time. GPS time is derived on the basis of computed transmit time. Table 1 summarizes the major sources of error in the technique described. Note that these errors are associated with a single measurement and do not include potential improvement from smoothing over multiple samples or combining measurements from multiple satellites. The reader is referred to (2) for a more complete discussion of the GPS Navigation Data Message and its contents.

TABLE 1

TTU ERROR BUDGET

Satellite Ephemeris	10 ns
Satellite Clock Drift	10 ns
Ionosphere/Troposphere	30 ns
User Location/Calibration	15 ns
Receiver Noise	20 ns
RSS Error	41 ns

2.2 Satellite Availability

NTS-2, the first of six NAVSTAR satellites planned for Phase I system deployment was launched in 1977. NDS-1 was launched in February 1978; NDS-2 was launched in May 1978, and are undergoing test and evaluation. Each satellite can be seen for an average of six hours per day anywhere in the world. Figure 3 illustrates expected times in view for the six satellite constellation from San Francisco. Note that at least one satellite is in view about 18 out of 24 hours each day.

3.0 RECEIVER DESCRIPTION

The Time Transfer Receiver is contained in a standard 19" wide rack mounted chassis. A remotable L-band preamp and omni-directional antenna are provided for external or roof top mounting.

3.1 RF Antenna/Preamp

The RF antenna/preamp assembly is designed to ensure a carrier-to-noise-density ratio for the L1 signal under worst case conditions of 37 dB/Hz. The antenna is a narrowband circularly polarized unit providing unit gain over the L1 frequency band everywhere above 20 degrees elevation. The preamplifier is of conventional design, preceded by a narrowband preselector, providing a noise figure of better than 5 dB and a gain of more than 40 dB. The entire assembly is designed for operation in an unprotected environment over a temperature range of -25oC to +65oC.

3.2 Time Transfer Receiver

The heart of the TTU is the Time Transfer Receiver (TTR). The TTR receives and tracks the PRN coded GPS signal at 1575 MHz, demodulates the carrier, and detects the 50 bps Navigation data which biphase modulates the carrier. Navigation data, code epochs, and receiver status are provided as outputs for processing or printing via an IEEE-488-1975 compatible interface bus every six seconds. Figure 4 illustrates the TTR front panel controls. The receiver may be operated manually from this panel or remotely via the processor interface. GPS time-of-week is continuously displayed in the center display. Figure 5 illustrates the major functional elements of the receiver.

3.3 Downconverter

The downconverter preselects, downconverts, amplifies, filters, and levels the input L-band signal. Single downconversion is used to an IF of 81.84 MHz where signal distribution and correlation is performed. All conversion signals are synthesized from the local frequency standard. Provisions are included for injection of test RF or IF signals from a NAVSTAR test transmitter such as STI Model 5001.

3.4 Code Tracking Loop

The code tracking loop is a conventional early-late delay-locked loop employing a single code channel which is switched between early and late code references. An early-late discriminator accepts the early and late codes, correlates the received and reference codes, and demultiplexes the switched error signals to provide a digital error signal proportional to the received code offset. The code loop filter and number-controlled-oscillator circuits which generate the corrected code clock are all digital designs which retain the phase noise purity of the reference oscillator so as to achieve a very narrow band tracking loop. The code generator generates any one of 37 Gold codes used for the GPS C/A signals, the proper code selectable by satellite ID.

3.5 Navigation Data Detection

This module accepts video 50 Hz navigation data from the carrier loop, synchronizes and detects the data and subframe synchronization, and generates data, timing and synchronization signals for the external processor. The received HOW word is also extracted for panel display.

3.6 Frequency Synthesizer

This module plus an internal 10 MHz frequency reference provides for generation of all internally used frequencies from an internal or external frequency standard.

This module provides control of receiver operating modes, front panel controls, local HOW generation and external control and status interfaces with the processor via the IEEE Interface Bus.

3.7 Data Processing Configuration

As indicated earlier, the data processing configuration required for the TTU will depend on the particular host user's application. For the development model of the TTU, a time-shared, 16-bit miniprocessor/CRT keyboard will be used. This configuration was selected to provide flexibility of test and evaluation, particularly with respect to evaluation of time-transfer algorithms. An extended precision time interval counter is included in the processor configuration to process the receiver generated code epochs measurements with 10 nanoseconds resolution.

The software provided with the TTU includes, in addition to standard vendor operating systems and library routines, nine applications peculiar modules which provide operator initiation, control, data acquisition, data reduction, display, and diagnostic capability, all at the CRT/keyboard which acts as the control station for the system. Figure 6 illustrates the processing sequence.

Upon initialization, the system verifies proper program loading, checks all Input/Output operations, and requests operator inputs for satellite ID, initial doppler offset and verification of proper system time. The receiver is then interrogated, and if status is valid, a navigation data frame is read in and checked for validity. When a valid frame of data is available, navigation data processing begins. This process produces an estimate of transmit-time once every six seconds which is then compared to the measured time-of-arrival as determined from the time interval between received and local code epochs. Consecutive differences are then smoothed over a two minute interval to produce refined estimates of both time and frequency error in the local user's clock. These data are displayed to the operator as illustrated in Figure 7.

4.0 CONCLUSIONS

A Time Transfer Unit capable of providing 100 nanoseconds world-wide time synchronization via RF signals transmitted by GPS NAVSTAR satellites has been described. Such units can provide more frequent and improved calibration of existing atomic clocks at remote locations, and upgrading of time and frequency accuracy available from conventional crystal frequency standards.

REFERENCES

1. Easton, R.L., Fisher, L.C., Hanson, D.W., Hellwig, H.W., Rueger, L.J., "Dissemination of Time and Frequency by Satellite," Proc. IEEE, Oct. 1976, Vol. 64, pp. 1482.

2. Van Dierendonck, A.J., Russell, S.S., Kopitske, E.R., Birnbaum, M., "The GPS Navigation Message," AIAA Guidance and Control Conference, San Diego, CA, August 1976.

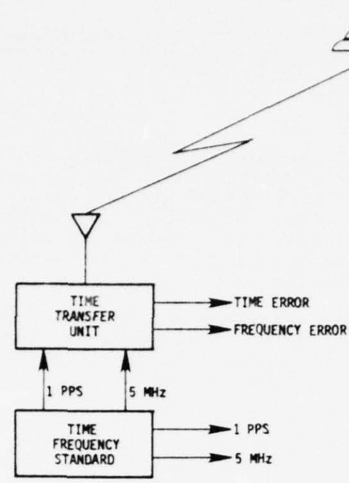


FIGURE 1 TIME TRANSFER UNIT

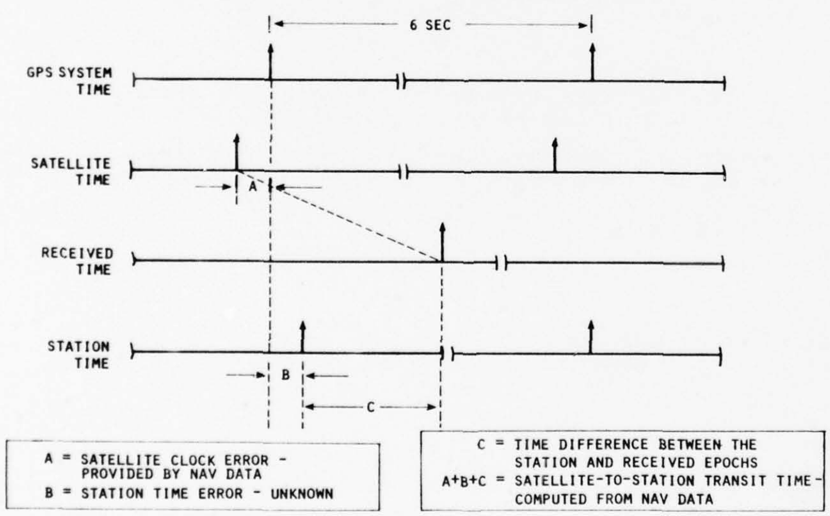


FIGURE 2 SYSTEM TIME RELATIONSHIPS

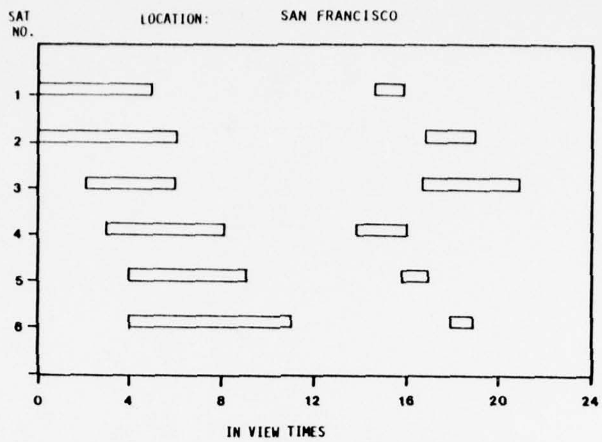


FIGURE 3 GPS PHASE I SATELLITE IN VIEW TIME

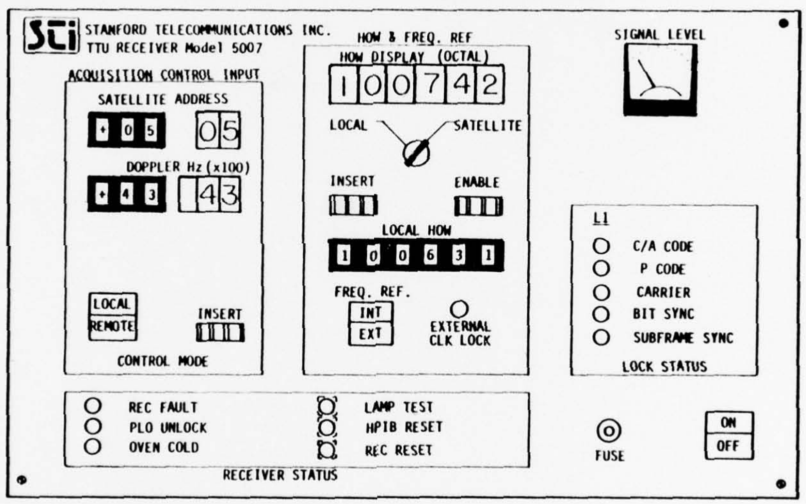


FIGURE 4 TIME TRANSFER RECEIVER CONTROLS AND INDICATORS

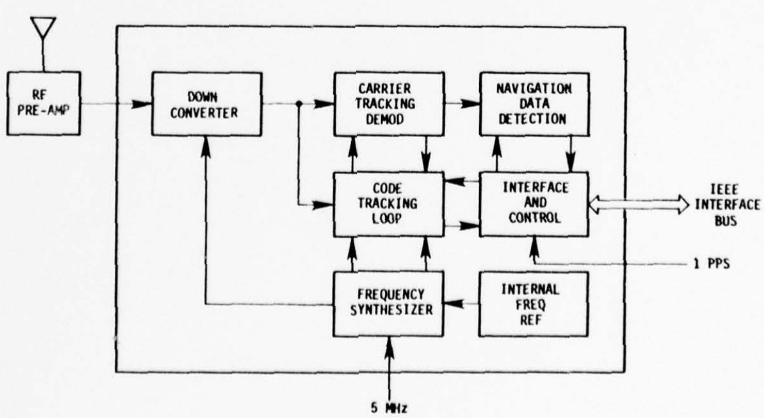


FIGURE 5 TIME TRANSFER RECEIVER

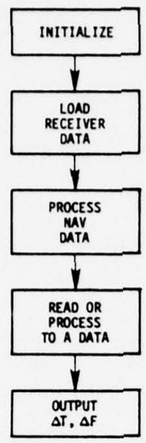


FIGURE 6 TTU PROCESSING SEQUENCE

GMT TIME DAY:HR:MI:SE	GPS TIME OF WEEK	TIME ERROR SECONDS	FREQ ERROR PARTS
123:12:23:40	12394	+5.11891E-07	+4.35179E-09
123:12:23:46	12395	+5.12510E-07	+4.26892E-09
123:12:23:52	12396	+5.16636E-07	+4.16312E-09
123:12:23:58	12397	+5.19073E-07	+4.06891E-09
123:12:24: 0	12398	+5.17294E-07	+4.00627E-09
123:12:24: 6	12399	+5.13459E-07	+3.93368E-09
123:12:24:12	12400	+5.14403E-07	+3.83309E-09
123:12:24:18	12401	+5.14089E-07	+3.72197E-09
123:12:24:24	12402	+5.15032E-07	+3.64947E-09
123:12:24:30	12403	+5.13194E-07	+3.57126E-09
123:12:24:36	12404	+5.18841E-07	+3.50480E-09
123:12:24:42	12405	+5.19879E-07	+3.42245E-09
123:12:24:48	12406	+5.20146E-07	+3.39225E-09
123:12:24:54	12407	+5.23184E-07	+3.33367E-09
123:12:25: 0	12408	+5.25976E-07	+3.27245E-09
123:12:25: 6	12409	+5.21148E-07	+3.23060E-09
123:12:25:12	12410	+5.20144E-07	+3.18878E-09
123:12:25:18	12411	+5.23869E-07	+3.10310E-09
123:12:25:24	12412	+5.23300E-07	+3.04279E-09
123:12:25:30	12413	+5.20568E-07	+3.01173E-09
123:12:25:36	12414	+5.21002E-07	+2.95748E-09
123:12:25:42	12415	+5.22653E-07	+2.89300E-09
123:12:25:48	12416	+5.19774E-07	+2.84795E-09
123:12:25:54	12417	+5.18780E-07	+2.81089E-09
123:12:26: 0	12418	+5.21168E-07	+2.75104E-09
123:12:26: 6	12419	+5.23914E-07	+2.70288E-09
123:12:26:12	12420	+5.19829E-07	+2.67354E-09
123:12:26:18	12421	+5.20393E-07	+2.64690E-09

FIGURE 7 TYPICAL TTU DISPLAY

EPHEMERIS AND CLOCK DETERMINATION IN GPS

J. H. Schaibly
General Dynamics
Electronics Division
P. O. Box 81127
San Diego, CA 92138
U.S.A.

E. L. Lasley
General Dynamics
Electronics Division
P. O. Box 81127
San Diego, CA 92138
U.S.A.

SUMMARY

The GPS Control Segment is responsible for providing accurate navigation data to each GPS satellite for subsequent broadcast to GPS users. This navigation data is based on satellite tracking data which is collected at the GPS Monitor Stations, processed by the GPS Master Control Station, and transmitted to the satellites by the GPS Upload Station. The present paper describes the data processing performed at the Master Control Station. Briefly, the tracking data is corrected for known deterministic effects and is used to generate optimal (i.e., minimum mean square) estimates of satellite ephemeris and clock behavior. Prior to satellite upload, the most recent satellite parameter estimates are propagated into the future to provide ephemeris and clock predictions which are then parameterized in a compact manner for the navigation message.

INTRODUCTION

The objective of the GPS Control Segment (CS) is to maintain the constellation of navigation Space Vehicles (SVs) in order to provide accurate, continuous, world-wide navigation support for GPS users. This maintenance consists of: a) The navigation support function; and b) The command and control functions. In Phase I of the GPS, b) is not a function of the CS per se, but has been delegated to the Air Force Satellite Control Facility. The present paper will be concerned only with the navigation support function.

The GPS satellites provide continuous RF transmissions from which users may determine their three-dimensional position. The transmitted data from each satellite includes a pseudorandom noise (PRN) signal from which the user may determine his distance from the satellite, and a low-rate (50 bit/second) data message which provides the user with the satellite's position (ephemeris). The data message also contains information about the on-board frequency standard (clock) from which the PRN code is derived.

The ephemeris/clock determination function is based on ranging measurements performed by four remote Monitor Stations (MSs)[1]. The MSs receive navigation and health data from each SV, perform pseudorange and integrated doppler measurements, and report all data to the Master Control Station (MCS). The MCS, in turn, processes these measurements using a Kalman estimator/predictor to predict the satellite ephemerides and clock behavior for some future span of time ($\sim$ 24 hours). These predictions are then encoded into the navigation message format[2] and transmitted via the Upload Station (ULS)[3] to the satellite on-board computer for subsequent transmission to the user.

The ephemeris determination technique used by the Control Segment is a two-step process making use of an off-line least-squares batch fit using approximately one week of measurement data to produce a reference ephemeris, and a first-order correction computed on-line by the MCS Kalman estimator using additional measurement data. The reference ephemeris provides both initial estimates of the satellite trajectory around which perturbations are computed by the estimator, and the time dependent state transition matrix (partial derivatives) also used in the estimator. The batch fit for the reference ephemeris is performed at the Naval Surface Weapons Center (NSWC/DL) at Dahlgren, Virginia using the CELEST program and measurement data transmitted daily to them from the MCS at Vandenberg AFB. A predicted reference ephemeris is subsequently sent from NSWC to the MCS weekly for the on-line processing of new measurements.* A time line for the two-step ephemeris determination process is shown in Figure 1 for one reference ephemeris span. This process is repeated every 7 days.

MCS Ephemeris Determination

The on-line Kalman estimator processing is summarized functionally in Figure 2, and is characterized by the following elements.

---

* A trajectory and partials integrator has also been included in the Phase I MCS for generating the reference ephemeris locally. This integrator can use initial conditions derived from Kalman estimates at a time t_e to produce a reference ephemeris for use in the filter for $t>t_e$. In this mode of operation, the ephemeris determination technique resembles an extended Kalman filter.

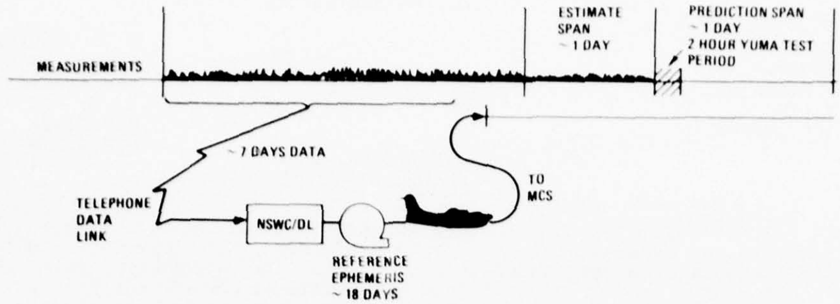


Figure 1. Time Line for Ephemeris Determination

Ranging Measurements — Ranging measurements (nominally every 6 seconds) are sent from each MS to the MCS for each satellite in view. The measurements include L-band signals at 2 frequencies, L_1 and L_2:

- L_1 pseudorange measurements
- L_1-L_2 pseudorange difference measurement
- Delta pseudorange measurement on L_1 over the measurement reporting interval (integrated doppler measurement).

Satellite and MS health and status data as well as meteorological data are also received.

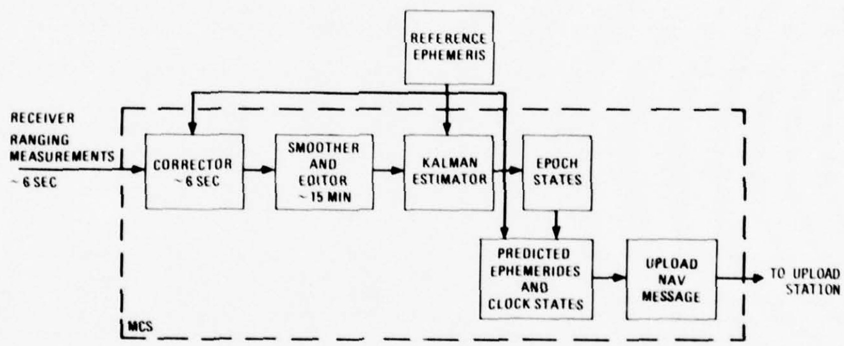


Figure 2. Functional Block Diagram of the MCS Ephemeris and Clock Determination Processor

Corrector — Corrections are made to each measurement to account for various known biases: Satellite and monitor station antenna phase center offsets; relativistic clock drifts; general relativistic frequency shift; special relativistic ray path aberration; and MS motions during signal transit time (because we have chosen to work with instantaneous geometric ranges rather than slant ranges). Tropospheric refraction corrections are computed using the model of Chao[4,5] in which temperature, humidity, and barometric pressure are parameters. The ionospheric delay correction uses the L_1-L_2 pseudorange difference measurement (the so-called "two-frequency model"). Time tags are also modified from MS receive time to SV send time. The range measurements are converted to residuals in the corrector by the subtraction of the satellite range as computed using the reference ephemeris. The corrected pseudorange measurement is then a measurement of

$$\left| \bar{R}_{SV} - \bar{R}_{MS} \right| - c\Delta t_s + c\Delta t_m + \frac{X}{\sin \varepsilon} - \left| \bar{R}_{Ref} - \bar{R}_{MS} \right| \qquad (1)$$

where each term is evaluated at the time of SV transmission. $\bar{R}_{SV}$ and $\bar{R}_{Ref}$ are the true and reference SV position vectors, respectively; $\bar{R}_{MS}$ is the MS location vector; Δt_s is the SV clock offset and Δt_m is the MS clock offset; and X is a residual (unmodeled) tropospheric delay term. The corrector also estimates the noise variance on the measurements.

Smoother — In the course of an iterative process of fitting and editing, the ranging data smoother generates a weighted least-squares polynomial fit to an edited subset of the corrected pseudorange and delta pseudorange data collected for each SV-MS pair. For a given pair, the polynomial

$$\hat{r}(t) = a_o + a_1 t + \ldots + a_n t^n \tag{2}$$

is then evaluated at the midpoint of the raw data span, t_M, and start and end times of the raw measurements t_S and t_E. The smoothed pseudorange measurement residual, $\hat{r}(t_M)$ and smoothed delta pseudorange residual,

$$\Delta\hat{r}\ (t_E,\ t_S) \triangleq \hat{r}\ (t_E) - \hat{r}\ (t_S) \tag{3}$$

are formed for use in the Kalman filter. In addition, the smoother estimates the corresponding noise (and the correlation coefficient) of the two smoothed measurements based on the noise estimates input from the corrector.

Kalman Estimator — The purpose of the Kalman estimator is to form revised estimates of the system state vector and the estimate error covariance matrix using the smoothed pseudorange and delta pseudorange measurement residuals. The Kalman filter estimation error covariance matrix is computed using the measurement noise estimates generated by the smoother and model dependent process noise parameters, and using the reference trajectories and partial derivative provided on the reference ephemeris. The filter implementation chosen is a Carlson square-root sequential filter with states carried at constant epoch (reference) times. A single computational cycle consists of a single time update,

$$S_o\ (t + \Delta t) = (S_o(t)\ S_o^T\ (t) + \Phi^{-1}\ (t)\ Q(t)\ \Phi^{-T}(t))^{1/2} \tag{4}$$

followed by a sequence of measured updates, each of the form

$$X_o^+ = X_o + K\ [z - H\ X_o] \tag{5}$$

$$S_o^+ = \left[(I - KH)\ S_o S_o^T \right]^{1/2} \tag{6}$$

$$K = \frac{S_o S_o^T\ H}{H\ S_o S_o^T\ H^T + R} \tag{7}$$

where:

K = Kalman gain matrix

S_o = Lower triangular square root of the estimate error covariance matrix

X_o = System epoch state vector

Φ = State transition matrix

Q = Process noise matrix

R = Measurement noise

H = Measurement matrix

The matrix square root indicated in equation (4) for S_o is evaluated using Cholesky factorization, while that in equation (6) uses the method of Carlson which ensures that S_o^+ is in lower triangular form.

The Kalman filter program also contains provisions for solution partitioning, reinitialization, redundant state recombinations, master clock designation changes, and special processing to handle such occasional or short term effects as SV momentum dumps, station keeping maneuvers, and MS receiver recalibrations.

The Phase I system state vector contains the following states:

$$\left.\begin{array}{l}\bullet\ \text{3 satellite position coordinates}\\ \bullet\ \text{3 satellite velocity coordinates}\end{array}\right\}\ \text{6 orbital elements, } P_o$$

- 3 solar pressure scaling parameters per satellite, SP_o

- 3 satellite clock states (phase, frequency, and frequency drift)
- 2 Monitor Station clock states (phase and frequency) for each nonmaster MS (by definition of system time, the master MS clock is perfect)
- 1 tropospheric residual bias state for each MS
- 3 polar wander states

For the Phase I system with 6 SVs and 4 MSs, 85 states are estimated.

The transition matrix, Φ *may be decomposed as shown in equation (8) for a partition* containing one SV.

$$\Phi(t) = \begin{bmatrix} \Phi_{E,E} & \Phi_{E,SP} & 0 & 0 & 0 \\ 0 & \Phi_{SP,SP} & 0 & 0 & 0 \\ 0 & 0 & \Phi_{SVc,SVc} & 0 & 0 \\ 0 & 0 & 0 & \Phi_{MS,MS} & 0 \\ 0 & 0 & 0 & 0 & \Phi_{PW,PW} \end{bmatrix} \quad (8)$$

where

$$\Phi_{E,E}(t,t_{oe}) = \left(\frac{\partial\bar{R}(t)}{\partial P_o}, \frac{\partial\bar{V}(t)}{\partial P_o}\right) \quad \text{(ephemeris)}$$

$$\Phi_{E,SP}(t,t_{oe}) = \left(\frac{\partial\bar{R}(t)}{\partial SP_o}, \frac{\partial\bar{V}(t)}{\partial SP_o}\right) \quad \text{(ephemeris - solar pressure)}$$

$$\Phi_{PW,PW}(t,t_{oe}) = \Phi_{SP,SP}(t,t_{oe}) = \Lambda \quad \text{(polar wander and solar pressure)}$$

$$\Phi_{SV_c,SV_c}(t,t_{oSV_c}) = \begin{pmatrix} e^{-a(\tau_{SV})} & 0 & 0 \\ \frac{1}{2}\left(1-e^{-a(\tau_{SV})}\right) & 1 & 0 \\ \frac{\tau_{SV}}{a} - \frac{1}{a^2}(1-e^{-a\tau_{SV}})\tau_{SV} & 1 \end{pmatrix} \quad \text{(SV clock frequency drift, frequency and phase)}$$

$$\tau_{SV} = t - t_{oSV_c}$$

$$\Phi_{MS,MS}(t,t_{oMS}) = \begin{pmatrix} 1 & \tau_{MS} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{(MS clock phase and frequency and troposphere)}$$

$$\tau_{MS} = t - t_{oMS}$$

where Λ is the identity matrix, and t_{oe}, $t_{o\,SV_c}$ and $t_{o\,MS}$ are the ephemeris, SV clock, and MS clock epoch times, respectively.

For partitions containing more than one SV, the Φ matrix is augmented to include the 12 X 12 combined ephemeris, solar pressure, and SV clock state matrix for the additional SVs. The MS and PW states are repeated in each partition.

The upload schedule requires that the convergence to the Kalman estimator be as rapid as possible. For this reason, the filter has a sophisticated process noise model including estimates of model error for solar radiation pressure, solar eclipse time uncertainty, geopotential uncertainty, and earth albedo uncertainty. Each of these effects enter with their approximate geometric dependence and correlation coefficients.

Ephemeris and SV Clock Prediction — The satellite clock states are propagated throughout the prediction span and a time polynomial representation is uploaded to the satellite. The SV reference ephemeris is corrected by the estimated system states by the equations

$$\bar{R}_{Pred}(t) = \bar{R}_{Ref}(t) + \frac{\partial \bar{R}(t)}{\partial P_o} P_o + \frac{\partial \bar{R}(t)}{\partial SP_o} SP_o$$

$$\bar{V}_{Pred}(t) = \bar{V}_{Ref}(t) + \frac{\partial \bar{V}(t)}{\partial P_o} P_o + \frac{\partial \bar{V}(t)}{\partial SP_o} SP_o$$

$\bar{R}_{Pred}$ and $\bar{V}_{Pred}$ are computed in earth-centered, earth-fixed coordinates using transformations containing the reference and state estimates for pole wander. The antenna phase center offset from the SV center of mass is added to the predicted position before fitting of the uploaded orbital element representation. An approximation to the user independent part of the relativistic frequency correction is also applied to the clock prediction polynomial representation before upload.

Simulated Performance

The ephemeris and clock estimation/prediction techniques described above have been evaluated in simulation for one SV. Using simulated smooth measurements from the four Phase I MSs in the so-called "Ephemeris End Around Check" the simulated SV orbit corresponded nominally to the Navigation Technology Statellite, NTS-II, and the simulated error sources included

- Tropospheric residual error
- Geopotential error
- Solar radiation reflectivity errors
- Receiver noise
- Random clock error (NTS-II cesium).
- Polar wander errors

The calculations spanned four weeks and simulated the estimation/prediction schedule expected in Phase I. The Aerospace Corporation generated simulated smoothed measurements both for NSWC to produce a reference ephemeris, and for GDE to generate predictions based on the Kalman filter estimation. The reference ephemeris for the first twelve days was differenced with the true SV trajectory at Aerospace, and the results are shown in Figure 3. The large in-track error is due to solar pressure modeling error. New reference trajectories were generated every 7 days.

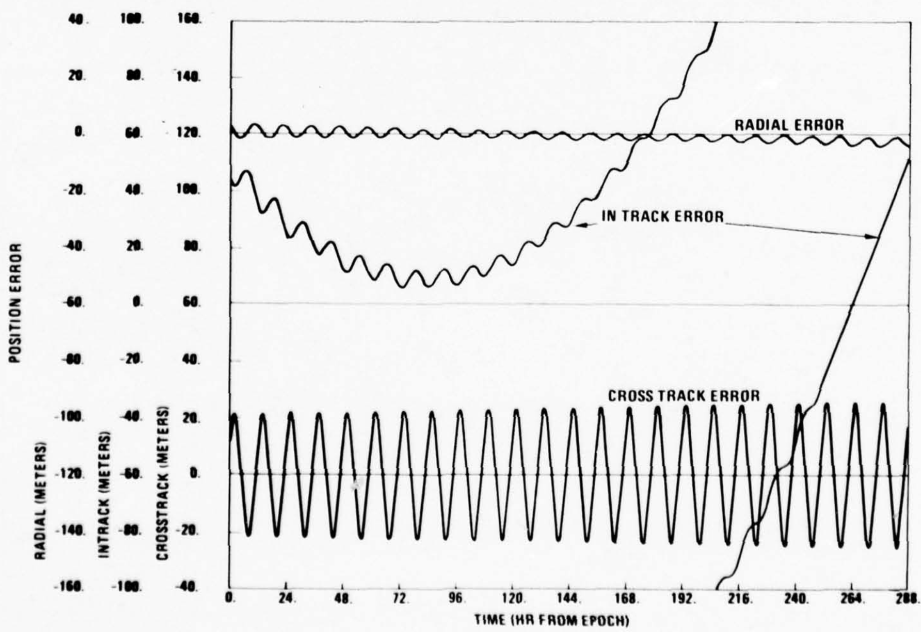


Figure 3. Reference Ephemeris Error — Week 1

The reference ephemeris was used in the Kalman filter for 28 days; the first 7 days without prediction to stabilize the filter, and the remaining days with 24-hour prediction spans following each 24 hours of estimation. The error between the Kalman filter predictions and the truth trajectory is shown in Figure 4 for each prediction span in the second week. It can be seen that the filter removed ~ 90% of the in-track error remaining in the reference ephemeris, and ~ 75% of the cross-track and radial error.

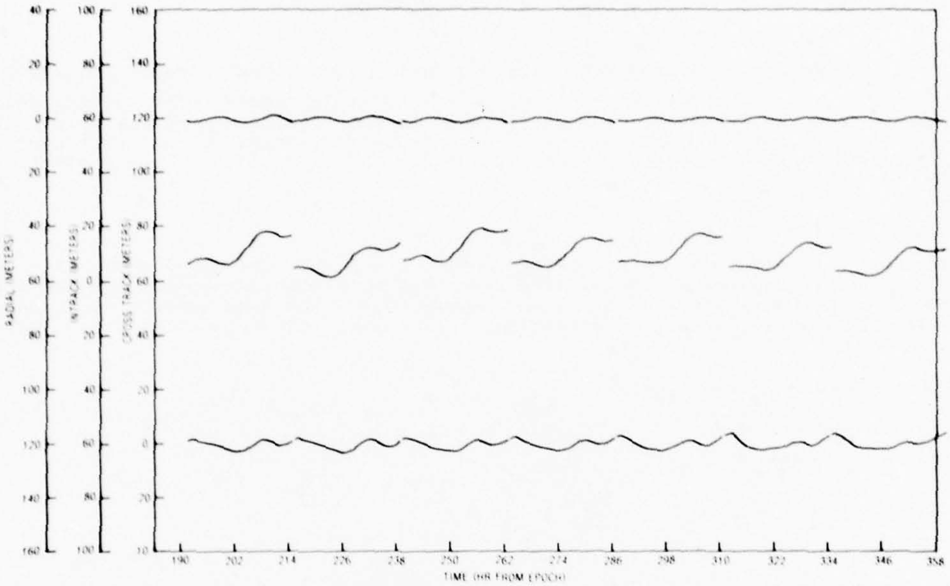


Figure 4. Error Between the Kalman Filter and Truth Trajectory Predictions during the Second Week

Figure 5 shows the SV clock prediction error for week two of the simulation. The nonuniform magnitude of the SV clock error reflects the random nature of atomic clock frequency.

The User Equivalent Range Error (UERE) is the pseudorange error for a particular location. It is composed of an ephemeris component and an SV clock component (as well as tropospheric delay, ionospheric delay, multipath, and user clock and receiver errors,

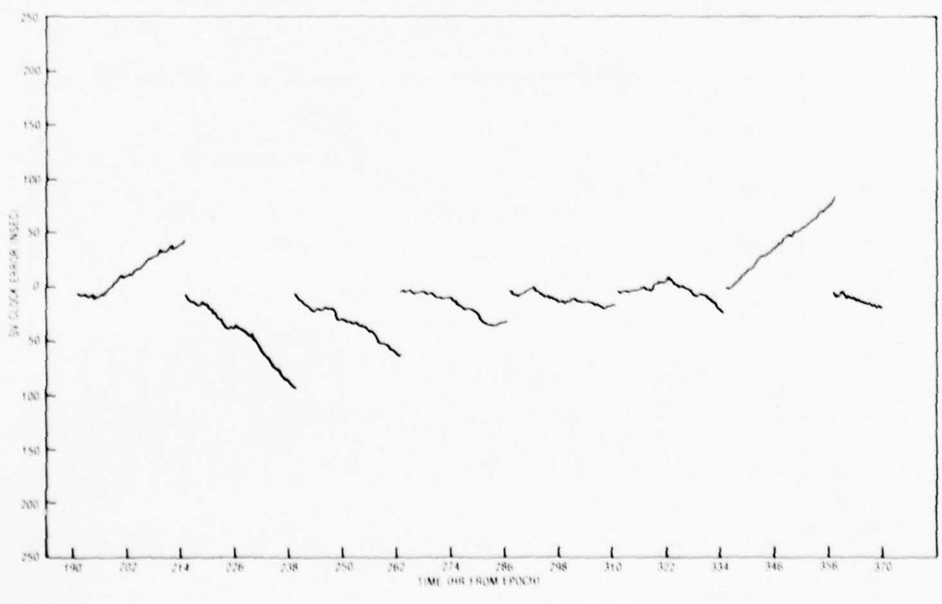


Figure 5. SV Clock Prediction Error for the Second Week

all of which we ignore in this discussion of CS performance). For a user at the Phase I Yuma Proving Ground test range, the simulated UERE is given below (RMS at 1, 2 and 3 hours after upload over the second, third and fourth weeks of simulation):

		1 hr	2 hr	3 hr
$UERE_{Ephemeris}$	=	1.46m	1.04m	.36m
$UERE_{SV\ clock}$	=	1.67m	1.91m	2.30m
$UERE_{Total}$	=	1.01m	1.59m	2.35m

The total UERE is smaller than the RSS of ephemeris and clock components due to error correlation.

The Phase I specification value for SV clock prediction is 9 ns (2.7m) after 2 hours, which is met in the simulations. The specification value for line of sight ephemeris error is 3.66 meters UERE after 24 hours. This is met according to this simulation which shows the average of the ephemeris error using the worst case formula:

$$UERE_{Ephemeris} \lesssim \epsilon_{radial} + .25 \sqrt{\epsilon^2_{in\text{-}track} + \epsilon^2_{cross\text{-}track}}$$

$$= 3.0 \text{ meters}$$

This overestimates the error as seen by the user since it is unlikely that he would choose 4 worst case SV's to navigate by.

Acknowledgements

The authors acknowledge the major role played by A. B. Bierman and W. A. Feess of the Aerospace Corporation in defining the End Around Check performance simulations and generating the measurement data. M. D. Harkins and B. Herman of NSWC generated the Reference Ephemerides for these simulations.

References

1. Harrington, R. L., W. A. Fabrizio, and E. Salinas, "Monitor Stations." pp___ this monograph.

2. Vandierendonck, A. J., S. S. Russell, and E. R. Kopitske, "The GPS Navigation Message," pp___ this monograph.

3. Harbour, W. H., and M. Rochette, "The GPS Upload Station," pp___ this monograph.

4. Chao, C. C., "A New Method to Predict Wet Zenith Range Correction from Surface Measurements," JPL Technical Report 32-1526, Vol XIV.

5. Chao, C. C., "The Tropospheric Calibration Model for Mariner Mars 1971," JPL Technical Report 32-1587.

TEXAS INSTRUMENTS PHASE I GPS USER EQUIPMENT

by

M.J. Borel, J.N. Damoulakis, D.R. Delzer, T.D. Fuchser,
J.H. Hinderer, C.R. Johnson, and D.J. Pinkos
TEXAS INSTRUMENTS INCORPORATED
13500 North Central Expressway
Post Office Box 226015
Dallas, Texas 75266 U.S.A.

SUMMARY

The use of standardized circuit modules enabled Texas Instruments to produce affordable Global Positioning System user equipment sets which satisfy the wide range of performance required by various applications, although the designs were restrained by the size and cost allowances of those applications. Such characteristics as time-to-first-fix, jamming resistance, and permitted host vehicle velocity were adjusted by changing the quantities of common modules that comprise the receivers and processors of the Manpack/Vehicular User Equipment, the High Dynamic User Equipment, and the Missile-Borne Receiver Set.

I. INTRODUCTION

Historically, sophisticated equipment was designed and fabricated to meet its specific requirements. Although such an approach may be the simplest to undertake, it provides only minimum production advantages to each user. Cost and risk for each configuration are high. In an attempt to offset some of these disadvantages, there is an inclination toward the development of three or four standard system configurations, one of which may be selected for a specific user application. Such an approach does afford some advantages as a result of the potential increase in production of each configuration, but, in most cases, it will result in compromises of performance and physical characteristics.

A more effective approach for achieving utility and affordability can present perhaps the more difficult initial design and development issues. This approach requires establishing user equipment commonality, not at the system level, but at the subsystem level. These building blocks, or common modules, can be used to satisfy every user equipment requirement and, consequently, provide the maximum production volume advantage. Because of the flexibility and commonality established at this level, performance and physical characteristics are not compromised for specific user requirements. Component technology improvements can be incorporated with minimum perturbations to other functional elements. The result is an approach that permits a concentration of effort in subsequent development phases aimed toward maximum cost improvements beneficial to all users. This approach is not encumbered with the difficulties of the "slight" variations of system performance that all too often require a major rework or redesign.

II. GLOBAL POSITIONING SYSTEM (GPS) USER EQUIPMENT IMPLEMENTATION

The design of common modules must be preceded by a careful definition of the functions to be performed by each module. This definition is evolved through an iterative evaluation process centered upon variables such as GPS equipment functions to be performed, current and projected component technology, performance and physical characteristics, and cost. Basically, this process follows the general steps listed below (illustrated in Figures 1 and 2):

Definition of functions to be performed by GPS user equipment

Separation of these functions into two categories: those sensitive to applications and those not sensitive

Subdivision of application-insensitive functions into elements grouped to provide maximum commonality over the identified user requirements

Design of common modules that provide the performance required of these elements.

The design of each common module was guided by imposing constraints defined by the requirement containing the worst case condition applicable to the module. For example, the design constraints for size, weight, and power may be driven by a man-portable requirement, while performance and environmental constraints may be driven by missile and satellite requirements. In addition, other constraints and controls were established that guided the common module life-cycle cost efforts.

The system concept and design resulting from this effort permit maximum commonality between various system designs through the use of different quantities of common modules to satisfy specific performance requirements. Also, because of the functional nature of the modules, improvements to accommodate performance requirements were incorporated with minimum impact. Texas Instruments GPS user equipment consists of three different system types: the High Dynamic User Equipment (HDUE) used in high-performance aircraft; the Missile-Borne Receiver Set (MBRS) used in the Minuteman missile test program; and the Manpack/Vehicular User Equipment (MVUE) used in manned vehicles such as trucks, tanks, and jeeps. A comparison of the major performance requirements of these systems is shown in Table 1, while a comparison of other characteristics is shown in Table 2. These systems are shown in Figures 3, 4, and 5.

Additional information regarding the three systems is presented following a discussion of the receiver section, the processor section, and the navigation filter section.

III. RECEIVER

A. Block Diagram

A receiver may be configured with one to five channels, depending on user dynamics and performance requirements. A block diagram showing one of these channels is presented in Figure 6. Each channel contains two or three narrowband modules and one

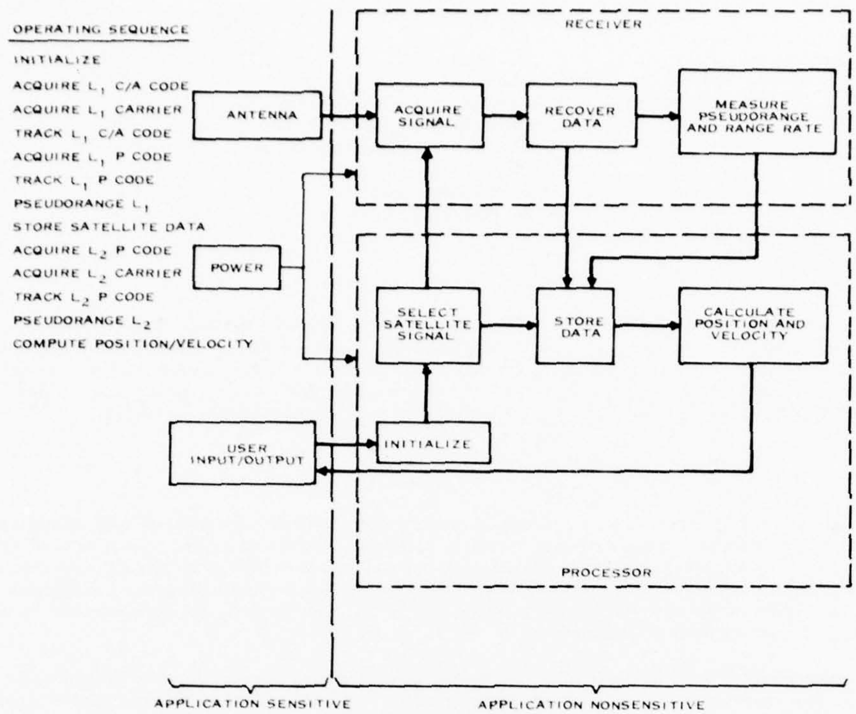


Figure 1. Functional System Operation

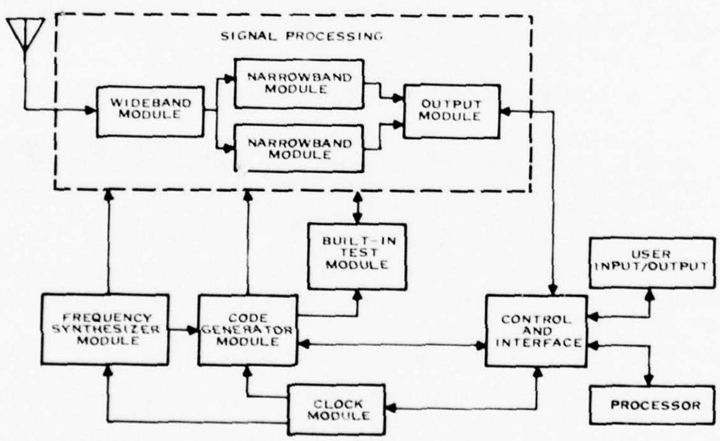


Figure 2. General Block Diagram

each wideband, output, code, and frequency synthesizer modules. Each receiver has a single master oscillator, clock module, built-in test module, and data processing unit, regardless of the number of channels. Both code and carrier loops are contained in each RF channel. The input signal of L_1 (154 f_0) and L_2 (120 f_0) is received at the antenna (f_0 is a reference frequency of 10.23 MHz). The code tracking is implemented in the narrowband modules (where correlation occurs), the output module and data processing unit (offset detection), and the code module (code generation and adjustment).

The carrier tracking loop is implemented in the wideband and narrowband modules (RF down-conversion and phase/frequency detection) and in the frequency synthesizer module (loop filter and VCO). The clock for the code generator module is tuned by the carrier tracking loop. Thus, the code loop tracking is aided by the carrier tracking loop.

B. Module Description

The input to the antenna is a doppler-offset carrier signal that is biphase modulated with code and data as shown in Figure 6. The antenna signal may be amplified in a low-noise preamplifier before reaching the wideband module. The wideband module amplifies the signal at L-band, down-converts to the first intermediate frequency (IF) of 18 f_0 (1 + V/C), and amplifies further in the IF. The IF bandwidth is 15 MHz to allow the code spectrum to pass. Some of the doppler is removed from the signal in all

TABLE 1. PHASE I GPS–REQUIREMENT COMPARISONS

Parameter	HDUE	MBRS	MVUE
SV carrier signal from antenna L_1 and L_2 (dBW)	–166 to –156	–176 to –166	–166 to –156
Code demodulation	P-code C/A code	P-code C/A code	P-code C/A code
User velocity (maximum) (meters/second)	1,100	7,620	25
User acceleration (maximum) (meters/second^2)	80	100	6
User jerk (maximum) (meters/second^3)	50	9	NA
Jamming levels, J/S (dB)			
Acquisition	24 (C/A)	NA	25 (C/A)
Track/lock	40 (P)	NA	40 (P)
Weak signal hold on	47 (P)	NA	NA
Time to first fix (seconds)	152	120	240
Pseudorange accuracy at (C/No)	1.5 meters (30 dB-Hz)	1.34 meters (25 dB-Hz) 2.4 meters (20 dB-Hz)	1.67 meters (30 dB-Hz)
Pseudorange rate accuracy (meters/second)	0.2	0.012	NA
Mean time between failures (hours)	500	Ps = 0.995	2,000
Temperature range	–20° to +55°C	+25° to +50°C	–40° to +70°C

TABLE 2. PHASE I GPS–HARDWARE COMPARISONS

Parameter	HDUE	MBRS	MVUE
No. of LRUs	4	1	1
No. of receiver channels	5	4	1
Size (feet^3)	3.5	1.6	0.75
Weight (pounds)	200	77	25
Power (watts)	529	208	45
Design-to-cost goal (1,000 units)	25K	NA	15K
Receiver modules/channel	6	7	6
Total common modules	65	50	12
Total unique units	13	13	6
Memory size (maximum capability)	68K	82K	48K
Memory size required	62.3K	62K	40K
Bits/word	16	16	16

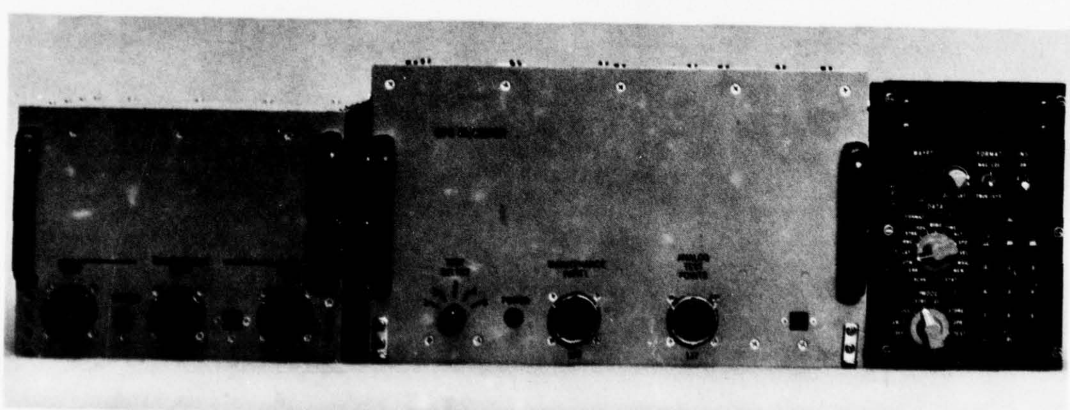


Figure 3. High Dynamic User Equipment

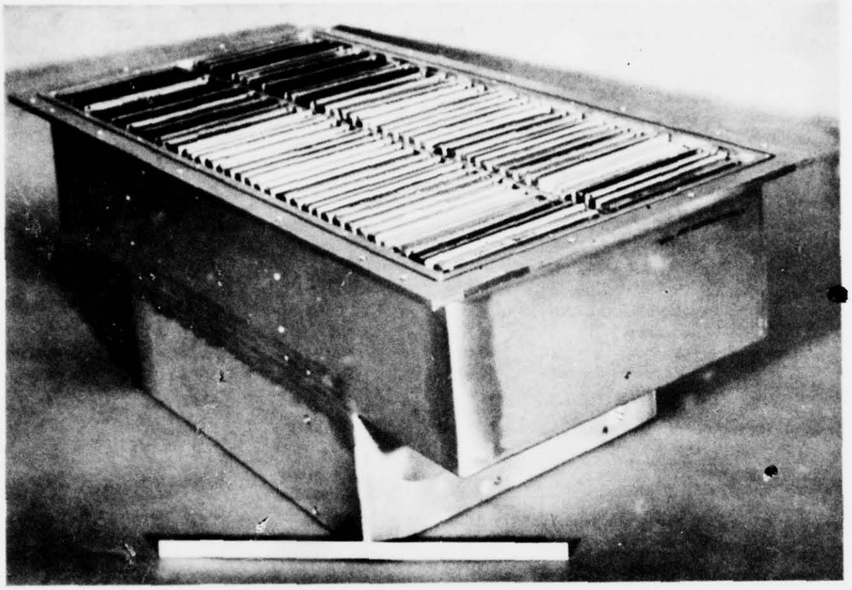


Figure 4. Missile-Borne Receiver Set

Figure 5. Manpack/Vehicular User Equipment

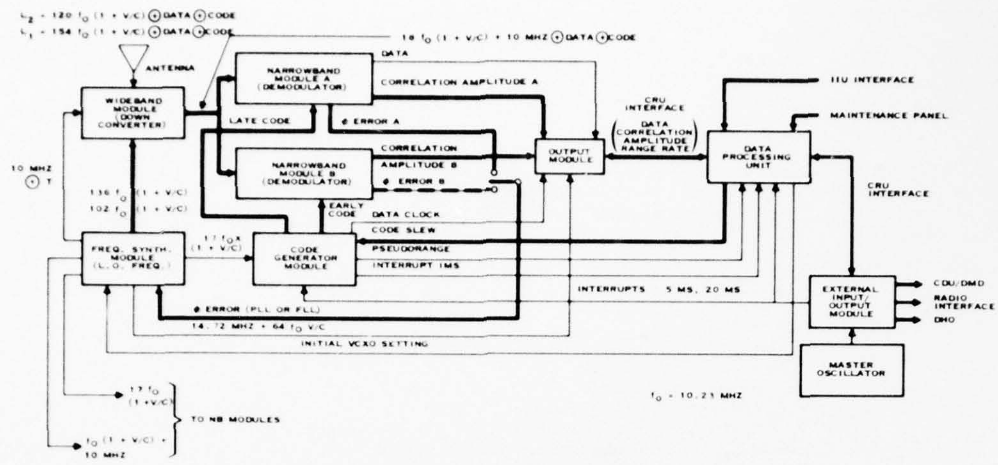


Figure 6. Manpack Receiver Functional Block Diagram

down-conversions in this receiver. When the carrier is locked, all local oscillator frequencies tune and track with the signal doppler offset in a coherent manner. The wideband module contains an AGC to control its gain and a pulse blanker to minimize the effects of pulse jamming. The output of this module offsets the IF by 10 MHz and introduces a receiver internal code (T-code) to enhance rejection of CW jammers. The wideband module output is split to feed either two or three narrowband modules.

The narrowband module has a second down-conversion and an IF at f_0 + 10 MHz. Final IF amplification and filtering occur before conversion to baseband in a Costas phase detector. The code is correlated or removed at this module's input, so the IF bandwidths are reduced to 50 kHz and eventually to 4 kHz to prevent noise power saturation in the IF amplifiers. Baseband circuitry in this module includes phase detection, frequency detection, envelope detection of the input signal, signal-to-noise ratio detection, and a second AGC circuit.

The phase or frequency detector output of the narrowband module is used to drive the carrier tracking network in the frequency synthesizer module. The network or loop filter output tunes a voltage-controlled crystal oscillator (VCXO) which tunes the frequency of all the local oscillators (LOs) used throughout the receiver. All LOs and the tracking network are in the synthesizer module. A digital loop filter and oscillator are used for these carrier tracking functions in the MBRS to accommodate tracking the higher doppler frequency and provide the more stringent accuracy in range rate measurements.

The code module generates the receiver's replica of the input signal code. Early, late, and prompt code versions are generated and output to the narrowband modules for code correlation and alignment. As described earlier, the code tracking is aided by the carrier tracking loop to remove the doppler. This is accomplished through 17 f_0 (1 + V/C) input to the code module used as the code clock after division by 17. In addition, the code can be slewed by division by 16 or 18 to center the code alignment. A code discriminator is formed in the data processing unit and is used to drive the code centering circuitry in the code module. The code module also contains the pseudorange measurement circuitry that measures the code state at a reference time mark.

The clock module uses a 10-MHz reference oscillator clock signal to provide timing marks needed in the rest of the receiver. Both hardware counters and processor interrupts are obtained from this module.

The output module serves as an interface between the RF hardware and the digital processor receiver control unit. The output module provides analog-to-digital (A/D) conversion of the envelope detectors in the narrowband modules and control decoding to set latches in the other hardware modules. The 50-Hz signal data are also detected in this module. Finally, a range rate counter is included in this module to count the frequency of the VCXO in the frequency synthesizer module. When the receiver is locked to the input signal, its output is the pseudorange rate used in navigation computations.

A separate built-in test (BIT) module provides a coded L_1 and L_2 test signal and an 18 f_0 test signal for testing the receiver. It allows calibration of time delays in multichannel systems and provides health checks for all the receiver channels.

C. Performance

The receiver provides pseudorange and pseudorange rate measurements for navigation computations. Table 3 shows the accuracy of these measurements for normal and jamming conditions.

IV. PROCESSOR

The various common processor system configurations are illustrated in Figure 7. All configurations use the same basic set of common processor modules (e.g., microprocessor module, memory modules, I/O modules). The single processor configuration uses fewer of each module type than the more sophisticated multiple processor configurations.

The microprocessor module (MPM) provides computational and functional control capability for the various system configurations. Each MPM interfaces with program memory modules (PMMs), data memory modules (DMMs), and a communication register interface module (CRIM) on its local memory bus. The components of the MPM, as shown in Figure 8, are the microprocessor unit, address decode logic, programmable read-only memory (PROM), PROM power switching circuitry, random-access memory (RAM), clock circuitry, and buffer logic.

The basic functional component of the MPM is the single-chip 16-bit I^2L SBP 9900 microprocessor unit. The 9900 is software compatible with the TI 990 minicomputer family. General operational characteristics the microprocessor unit exhibits are:

TABLE 3. RECEIVER ACCURACIES

	Pseudorange Accuracy (meters)	Pseudorange Rate Accuracy (meters/second)
Phase lock		
Normal	1.67	0.046
Jamming (J/S = 42 dB)	1.70	0.139
Frequency lock		
Normal	1.70	0.25
Jamming (J/S = 45 dB)	1.87	4.7

16-bit instruction word

3-MHz basic clock

Memory-to-memory architecture

Memory address capability for up to 32,768 sixteen-bit words or 65,536 eight-bit bytes

Separate memory, I/O, and interrupt bus structure

Use of 16 workspace registers in memory

Up to 16 prioritized interrupts

Instruction-driven communication register unit (CRU) and direct memory address (DMA) I/O capability.

The MPM address decode logic performs memory address recognition and decode for the memory (read-only or read/write) contained within the MPM. The PROM on the MPM provides the microprocessor unit with 512 words of nonvolatile storage for program instruction and data constants. The PROM switching circuitry minimizes MPM power by disabling the power source from all MPM PROM devices not being addressed. The RAM on the MPM provides the microprocessor unit with 256 words of high-speed read/write memory for allocation as workspace memory. The MPM clock circuitry is provided for use as a system clock at the user's option. The buffer logic on the MPM provides the necessary buffering for the microprocessor unit memory bus and CRU bus signals.

The MPM memory bus provides the mechanism for information transfer between the MPM and the memory for instruction fetch operations and storage/data retrieval operations. The instruction-driven CRU bus, along with the microprocessor unit DMA I/O feature, provides the MPM with input/output capabilities.

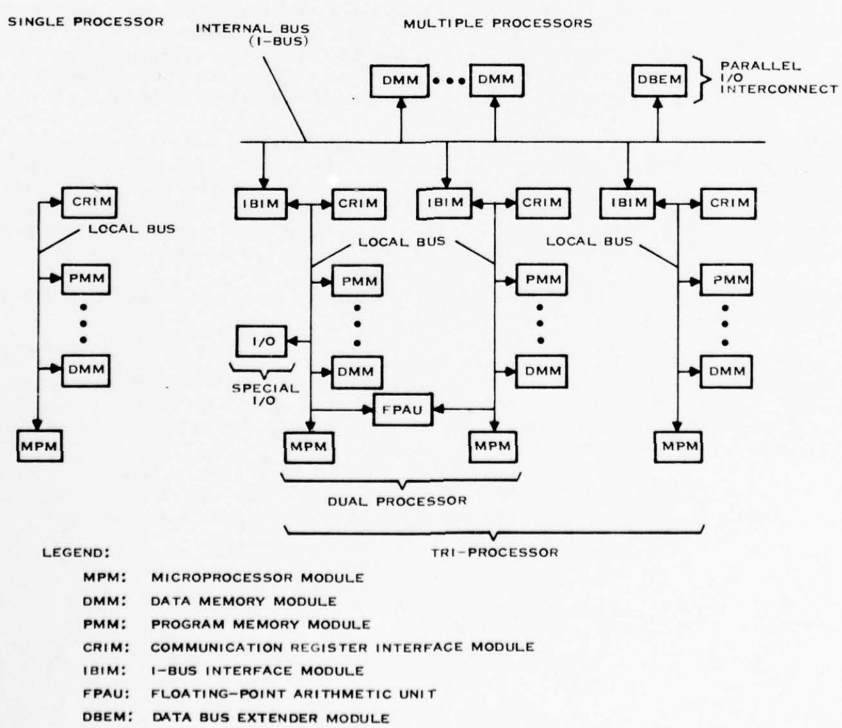


Figure 7. Common Processor Module Family

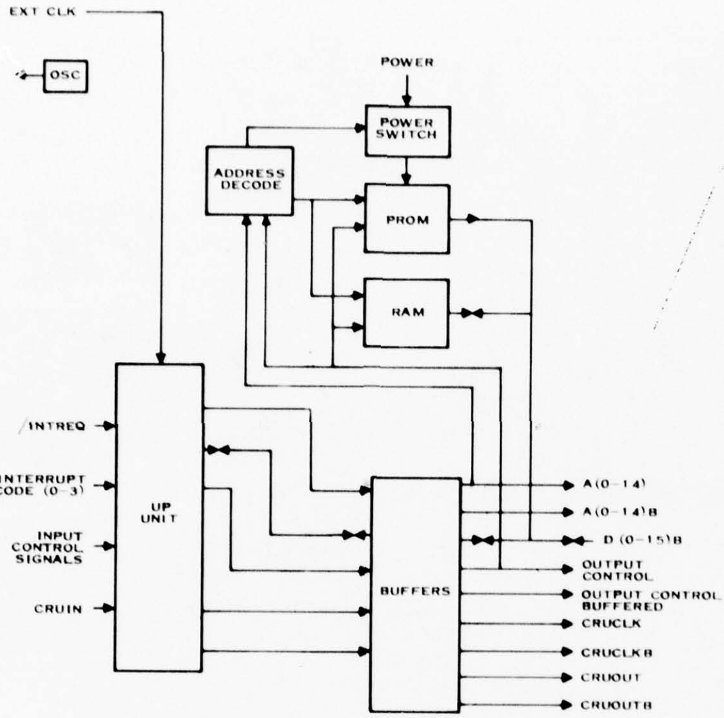


Figure 8. Microprocessor Module (MPM) Block Diagram

Each DMM provides the system with 4,096 words of random-access read/write memory for temporary storage of 17-bit data words (16 bits for data and 1 bit for parity). Each DMM contains a single-port data and address bus compatible with the MPM local memory bus.

Each PMM provides the system with nonvolatile PROM for program instructions and data constants. Each module contains a maximum of 16,384 words of 16 bits each. The PMM single-port data and address bus is compatible with the MPM local memory bus. The memory bus signal pinouts of the PMM and DMM are assigned so that a PMM can be inserted in any DMM location and replace up to four DMMs. The PMM minimizes power use by disabling the power source from any PROM devices not being addressed.

The CRIM consists of:

- CRU decoder and buffers
- Interrupt circuitry
- Parity generator/checker
- Reset circuitry.

The CRIM decodes bits 3 through 5 of the address bus into eight register select lines for use by any CRU device. The other address and CRU control lines are also buffered for use by CRU devices. The interrupt circuitry synchronizes the interrupt stimuli, provides interrupt masking/clearing capability, and generates the interrupt request and code signals for the MPM. The parity generator/checker performs parity checks on read operations and generates the parity bit for storage during write operations to read/write memory. Parity errors are signaled to the user as interrupts. The reset circuitry receives the various system reset stimuli (e.g., from power supply) and provides corresponding reset signals to the other processor modules.

In the multiple processor configurations (Figure 7), communication between the processors is accomplished via the internal bus (I-bus), a high-speed, 16-bit parallel data bus. The transfer of data between processors is accomplished through an intermediate DMM interfaced to the I-bus. Each MPM interfaces to the I-bus through its associated I-bus interface module (IBIM). The IBIM controls access onto the I-bus and provides for the bidirectional passage of data between the MPM and the I-bus.

In multiple processor configurations, a floating-poing arithmetic unit (FPAU) can be included to provide floating-point arithmetic and conversion capability. The FPAU is a dual-port, high-speed, auxiliary arithmetic unit that performs single and double precision arithmetic operations. The FPAU can be interfaced to a maximum of two processors via their local memory buses. In systems where the FPAU is interfaced to two processors, a software calling sequence has been established to avoid usage conflicts.

Additional input/output capability is provided by the data bus extender module (DBEM) which extends the I-bus to external devices. Other special I/O interface modules may be added on any local bus to satisfy system-unique I/O requirements.

The various processor modules and processor configurations are shown in Figure 9. Each module is implemented on a single 4.7- by 5.6-inch multilayer printed wiring board (PWB). The exception is the FPAU, which consists of seven multilayer boards.

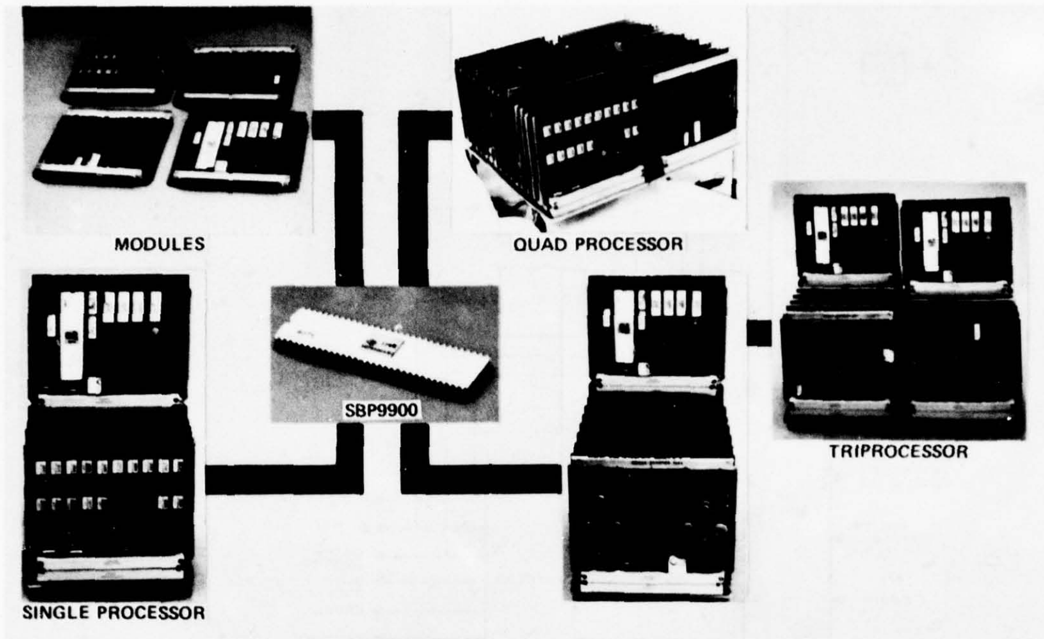


Figure 9. Processor Common Module Family

Interconnection of the various modules to obtain the different processor configurations is accomplished via multilayer mother boards.

The microprocessor unit on each MPM executes a full minicomputer instruction set containing 69 instructions. The arithmetic instructions include add, subtract, compare, negate, absolute value, increment, decrement, shift left/right arithmetic, clear, multiply, and divide. A variety of logical instructions is also provided. Included are set 1's/0's corresponding, compare 1's/0's corresponding, exclusive OR, invert, AND, OR, and shift logical/circular. The set of program control instructions includes jump (13 conditions), load, store, move, swap byte, branch, and return (from interrupt/subroutine).

Seven addressing modes are available for operand derivation. Instruction execution times range from 8 (add) to 124 clocks (divide). Most instructions require 10 to 20 clocks to execute, with variations dependent on the instruction function and the specified operand derivation cycle. Table 4 lists the various processor instruction execution times assuming a 3-MHz system clock.

V. NAVIGATION FILTER

The current user sets (HDUE, MBRS, and MVUE) employ a navigation filter configuration that optimizes the processing of pseudorange and pseudorange rate measurements. The HDUE and MBRS filters process the measurements simultaneously, while the MVUE filter processes the measurements sequentially.

For either type of processing, the system is as shown schematically in Figure 10. The system uses two computational loops: the fast loop (FL) and the slow loop (SL). The FL consists of the propagation (PR) task and the measurement-incorporation (MI) task. The SL consists of the optimal-filter (OF) task. The operations performed by each task are as follows:

Propagation (PR) task—propagates the dynamic equations forward in time. For this, assume that $\underline{y}$ is the state of the system. Underscored quantities are vectors; e.g., $\underline{y}$ represents vector with components $(y_1, y_2, \ldots, y_n)$. This task performs the operation:

$$\underline{y}_i = \Phi(t_i, t_{i-1})\, \tilde{\underline{y}}_{i-1} \qquad (1)$$

where $\underline{y}_i$ denotes the state vector of the system at the time t_i, $\tilde{\underline{y}}_{i-1}$ the state of the system at t_{i-1}, and $\Phi(t_i, t_{i-1})$ the system transition matrix from t_{i-1} to t_i.

Measurement-incorporation (MI) task—incorporates the receivable measurements in the system to correct the existing propagated state $\underline{y}$. This task performs the operation:

$$\tilde{\underline{y}}_i = \underline{y}_i + K_i\, \Delta z_i \qquad (2)$$

where K_i is a set of existing gains at t_i and Δz_i the measurement residual at t_i.

Optimal filter (OF) task—simultaneous processing (HDUE, MBRS) uses two stages, both of which employ an optimal Kalman filter. Sequential processing (MVUE) uses only one stage and also employs a Kalman filter.

In the HDUE and MBRS systems, the first stage, Stage I, propagates and updates the system error state covariance based on the time interval Δt; while the second stage, Stage II, propagates and updates the system error state covariance based on the time interval $\Delta\tau$. The output of the OF is a set of optimal gains K_i used by the MI task, and the system error covariance matrix P_i.

In the MVUE system, Stage I, the only stage, propagates and updates the system error state covariance based on the time interval ($\Delta t + \Delta\tau$). Similarly, the output of OF is a set of optimal gains K_i, used by the MI task, and the system error state covariance matrix P_i.

Basically, the system receives observable data, consisting of the vectors $\underline{p}$ and $\Delta\underline{p}$, at Δt intervals. The data are processed as they are received by the FL mechanization. The SL mechanization operates in parallel with the FL mechanization, in real time, computing the appropriate gains for the system and updating the system statistics. The execution of SL occurs at ($\Delta t + \Delta\tau$) time intervals.

A. Detailed System Description

Figure 10 describes the tasks relative to the processor time line. The process starts at time t_0 and executes one cycle of the OF, which takes the process through time t_*. Measurement data are received at regularly spaced time intervals Δt. These measurements are denoted at each instant as $\underline{p}_M$ and $\Delta\underline{p}_M$. The filter gains used during the execution of OF, denoted by K_0, correspond to the gains used to incorporate the measurements received at the particular times.

The process has available, at time t_0, an estimate of existing state $\underline{y}_0$. The process first propagates this state through the interval Δt between t_0 and t_1 to produce a predicted state $\underline{y}_1$ at time t_1. This extrapolation is done by the PR task. The predicted state $\underline{y}_1$ is then corrected by incorporating the data values received at t_1 using the existing set of gains K_0 in the task labeled MI. The resulting corrected state $\tilde{\underline{y}}_1$ is then the available system state at time t_1. This state will be propagated and corrected by another cycle of PR and MI tasks to compute the state $\tilde{\underline{y}}_2$ at t_2. The sequence of these events (an MI task followed by a PR task) constitutes the fast loop cycle. The FL cycle is repeated every Δt seconds throughout the execution of the OF cycle.

TABLE 4. INSTRUCTION TIMES

Instruction Type	Execution Times (microseconds) Minimum	Maximum
Arithmetic		
Add/subtract	4.67	10.00
Multiply	17.33	20.00
Divide	30.67	44.00
Compare	4.67	10.00
Shift (left/right arithmetic)	4.67	17.33
Absolute value	4.00	7.33
Increment/decrement	3.33	6.00
Clear	3.33	6.00
Logical		
Set 1's/0's corresponding	4.67	10.00
Compare 1's/0's corresponding	4.67	7.33
Exclusive OR	4.67	7.33
OR/AND (immediate)	4.67	4.67
Shift (logical/circular)	4.67	17.33
Swap bytes	3.33	6.00
Program control		
Move	4.67	10.00
Jump	3.33	3.33
Branch	2.67	5.33
Return	4.67	4.67
Load	4.00	4.00
Store	2.67	2.67
Input/output		
Single bit	4.00	4.00
Multiple bit	7.33	22.67

In each of the fast loops, the system is using the most recent set of gains available to incorporate the measurements. However, processor execution load prevents the system from computing a new set of optimal gains for incorporating each received measurement in real time. Instead, the process executes two filtering subtasks, in parallel with the execution of the fast loops, to create a new set of gains. The execution of these two filtering subtasks constitutes the slow loop cycle. Each slow loop is computationally equivalent to the optimal filter task.

The combination of the two subtasks which produces the new set of gains constitutes the Stage I filter followed by the Stage II filter. The two stages are serially executed.

Initially, the Stage I filter uses the state $\underline{z}_i = \underline{y}_0$ at time t_0, the receiver data at t_0, and the predicted covariance estimate M_0. Thereafter, the Stage I filter uses as starting values the output of the previous Stage II subtask and the last (current) output of the FL cycle. The Stage I filter takes the above information to compute a Stage I set of gains and updates the state covariance matrix and the system states. The resulting covariance P_1 and state vector $\underline{x}_1$ are based upon the user-to-transmitter geometry at time t_0 (or, in general, at time t_*). During the final step of the Stage I filter cycle, the system error state covariance and state are propagated forward at time Δt, producing the matrix P_1 and the state vector $\underline{x}_1$ which are valid for time t_1.

Upon completion of Stage I, the Stage II filter subtask begins. Input data for Stage II is the Stage I output P_1 and $\underline{x}_1$, plus the observed data $\underline{p}_M$ and $\Delta\underline{p}_M$, valid for time t_1. The filter processes this information to produce a Stage II set of gains and to update the state covariance. Note that, during the next SL filter cycle, the FL filter cycle will use these same optimal gains to incorporate measurements while yet another set of optimal gains is being created. The Stage II subtask is completed with the forward propagation of the state covariance over a time $\Delta\tau$; thus, the state covariance P_2, valid for time t_*, is generated. This last value of the covariance P_2 becomes the *a priori* covariance for the next SL filter cycle. The existing filter state is not propagated to time t_*. Instead, the *a priori* state vector $\underline{z}_*$ is set equal to $\underline{y}_*$ which is the output from the FL cycle at time $t_* - \Delta t$.

Thus, the fast loop consists of one execution of the MI task followed by one execution of the PR task using the best estimate of gains. The slow loop consists of one execution of the Stage I subtask followed by one execution of the Stage II subtask to produce the final optimal gains and system statistics. Stage I operates on the measurement and predicted state at time t_0 (or t_*) and the covariance estimate from the last Stage II to produce an updated covariance valid at time t_1. Stage II operates on the measurements at time t_1, the predicted state at t_1, and the covariance derived from the execution of the preceding Stage I.

The result of this process is an estimate of current system state based on more measurements than could be used if optimal filtering were used on-line. The estimate is more accurate than that obtainable from a single measurement incorporation during an optimal filter loop, or from simple data preaveraging.

For the MVUE sequential processing, the same philosophy as outlined above is followed with the exception that the Stage II filtering procedure is not present. At each fast loop, the measurements from only one satellite are processed. Thus, for four satellites, the fast loop is activated four times before completing one execution of the slow loop. During the execution of each slow loop, the gains and the error state covariance for only one satellite are updated. Therefore, it takes four executions of the slow loop to update the gains and covariance for all four satellites.

TABLE 5. TYPICAL STATISTICAL VALUES OF THE ERRORS FOR SIMULTANEOUS PROCESSING

Quantity	Average RSS	Spherical Error Probable	Circular Error Probable	Probable Error
Resultant position error in earth-fixed coordinates (meters)	11.74	10.36	–	–
Resultant velocity error in earth-fixed coordinates (meters/second)	4.75	3.47	–	–
Horizontal position error (meters)	–	–	10.50	–
Horizontal velocity error (meters/second)	–	–	2.89	–
Error in range bias (meters)	–	–	–	1.70
Error in range-rate bias (meters/second)	–	–	–	0.02
Error in altitude (meters)	–	–	–	2.90
Error in vertical velocity (meters/second)	–	–	–	1.15

TABLE 6. TYPICAL STATISTICAL VALUES OF THE ERRORS FOR SEQUENTIAL PROCESSING

Quantity	Average RSS	Spherical Error Probable	Circular Error Probable	Probable Error
Resultant position error in earth-fixed coordinates (meters)	20.72	–	–	–
Resultant velocity error in earth-fixed coordinates (meters/second)	1.49	–	–	–
Horizontal position error (meters)	–	–	13.08	–
Horizontal velocity error (meters/second)	–	–	1.18	–
Error in range bias (meters)	11.71	–	–	–
Error in range-rate bias (meters/second)	0.231	–	–	–
Error in altitude (meters)	–	–	–	9.99
Error in vertical velocity (meters/second)	–	–	–	0.634

TABLE 7. MAJOR MVUE CHARACTERISTICS

Size	23,000 cubic centimeters
Weight	15 kilograms
Major system outputs:	
Present position (MGRS or latitude/longitude)	
Range and bearing to any of 8 waypoints	
Time of day (day of week, hours-minutes-seconds)	
Battery life (with NiCad BB590)	
Continuous mode	2 hours
Periodic (4 fixes/hour)	6 hours
Stationary user CEP	15 meters
Dynamic user CEP (25 meters/second maximum velocity, 6 meters/second^2 acceleration)	50 meters
Receiver contribution to range accuracy (1σ)	1.64 meters
C/No thresholds (including jamming effects)	
Acquisition	34 dB-Hz
Accurate navigation	33 dB-Hz
Carrier/code track	30 dB-Hz

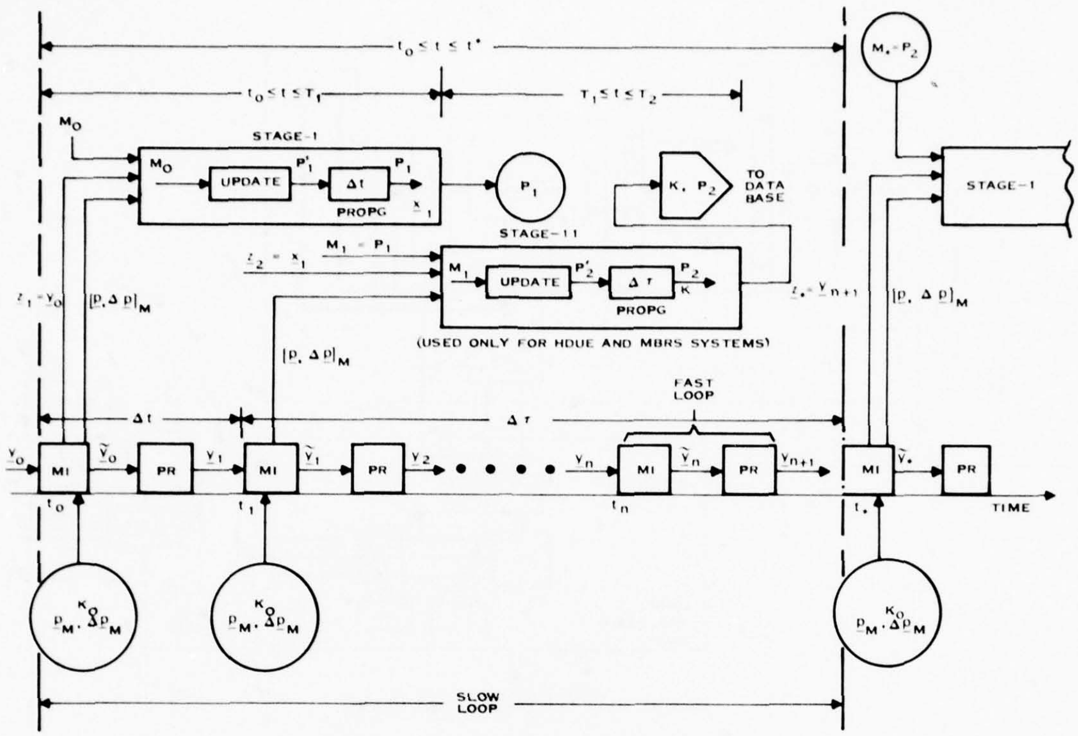


Figure 10. Schematic Representation of the Navigation Filter Structure

B. Simulation Results

Tables 5 and 6 present typical statistics of the errors resulting from implementation of the previously discussed filtering schemes, both simultaneous and sequential. Table 5 presents the results of an F-4 aircraft trajectory for 260 seconds of flight time. The satellite constellation is typical for GPS. As noted, these numbers may vary with the changes in the satellite constellation, trajectory, and measurement errors. Table 6 presents the results for a truck path at Yuma Proving Grounds. The simulation time is about 3,000 seconds. The satellite constellation and errors are the same as in Table 5. Again, it is noted that these results may change as the various error parameters and satellite constellations change. Details on the above simulation results are given in References 1 through 3.

VI. MVUE

The MVUE set consists of a single-channel sequencing receiver with its associated control/navigation data processor. A block diagram of this set is shown in Figure 11. The MVUE major characteristics are listed in Table 7.

The MVUE set sequences between satellites every 2 seconds. The major space vehicle (SV) selection, acquisition, and tracking processes are shown in Table 8. The MVUE tracks both L_1 and L_2 satellite frequencies and uses both C/A and P codes during the acquisition process. Navigation calculations are made using an eight-state Kalman filter.

Input initialization data (approximate position and time) are entered by means of a hand-held control/display unit (CDU) calculator-type keyboard. Output data are displayed on the LED display. Various outputs available from the MVUE set are shown in Table 9.

A number of advanced component technologies are used in the MVUE set to reduce size, weight, and power. These include miniature volute antenna, surface acoustic wave (SAW) bandpass filters, custom large-scale integration (LSI) digital circuits for code generation, an integrated injection logic (I^2L) 16-bit microprocessor, and hybrid RF circuits.

The MVUE set can operate from rechargeable or primary batteries (24 Vdc) or vehicle generator power. In the vehicle operation mode, a separate antenna/preamplifier is provided in a vehicle installation kit.

VII. HDUE

The HDUE set consists of four major line replaceable units (LRUs). They are the receiver, data processor, ac/dc converter and CDU. The receiver LRU is composed of five continuous tracking receiver channels and a receiver controller/processor. An existing antenna/preamplifier will be used for tests. The data processor LRU consists of the master state controller and the navigation processor. All GPS controller/processors are designed around the SBP 9900 microprocessor. The CDU has a calculator-type keyboard and an incandescent light alphanumeric display. A block diagram of this set is shown in Figure 12. Major characteristics of the HDUE are presented in Table 10. The HDUE major modes of operation are shown in Figure 13.

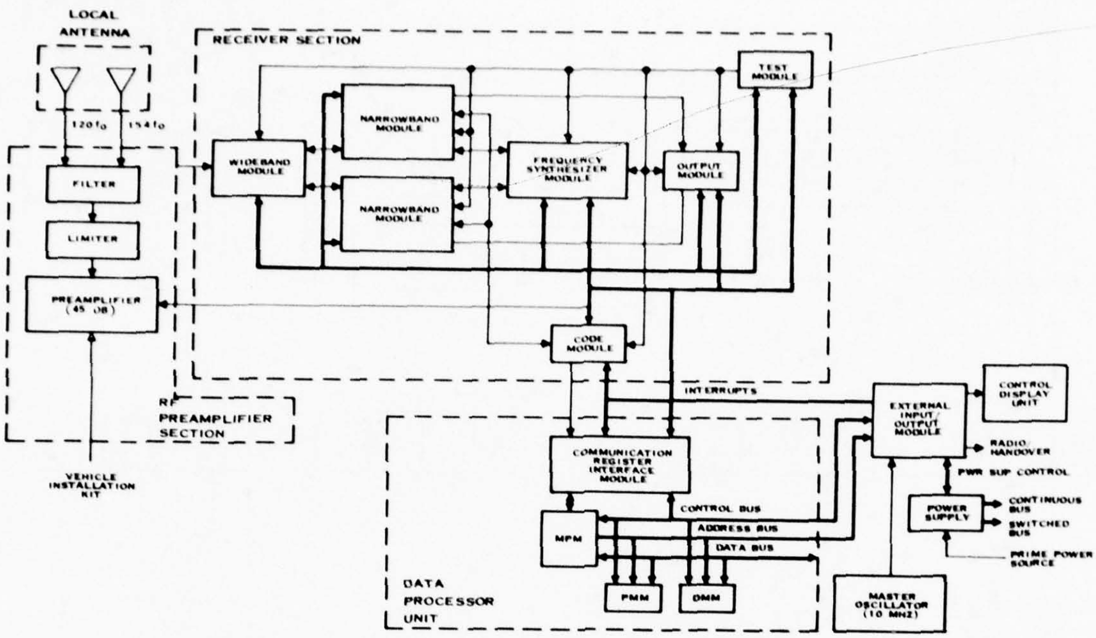


Figure 11. MVUE LRU Functional Block Diagram

TABLE 8. SV SELECTION, ACQUISITION, AND TRACKING PROCESSES

- SV Selection
  - Check every 2 minutes
  - Track longer to minimize changes
  - Fit parabola to orbit for navigation
- First SV Acquisition
  - Highest SV first
  - 25 kilometers, 30 seconds, and almanac require two full doppler bins
  - C/A "sequential" search
  - Code center
  - Phase lock
  - BIT synchronization
  - Frame and HOW synchronization
  - Reset time to 0.1 millisecond
  - C/A to P while still recovering data
  - Recover clock and ephemeris data
  - Measure pseudorange and delta range for navigation
- Subsequent SV acquisition
  - 25 kilometers, 0.1 millisecond, and almanac require one doppler bin
  - 600 C/A chip "sequential" search
  - 12 seconds during subframes 4 and 5
  - C/A to P while still recovering data
  - Recover clock and ephemeris data
  - Measure pseudorange and delta range for navigation
- Tracking
  - C/A reacquisition to improve aiding
  - P reacquisition every $\Delta\tau$
  - Track in deterministic sequence
    - Carrier/code track
      - Measurements for navigation
    - Accurate navigation
      - Ionospheric correction
      - Data refresh
      - New SV acquisition

TABLE 9. MVUE OUTPUTS

Position Output	Additional Outputs
Coordinates	Time
Grid	Altitude
Geographical	CEP
Timing Options	PE
Automatic mode	Number of SV
Manual mode	Waypoint
Time-to-first-fix mode	Position
Cold start	Range and bearing
Warm start	Audio digital (digital message device)
Direct handover	Instrumentation
Periodic mode	Direct handover

TABLE 10. HDUE MAJOR CHARACTERISTICS

Range measurement accuracy (J/S = 30 – 40 dB)	
Coarse (C/A code)	15 meters (1σ)
Fine (P code)	1.5 meters (1σ)
Resulting position accuracy	
Horizontal	10 meters (CEP)
Vertical	10 meters (PE)
Intrinsic maintainability design	
LRU replacement	5 minutes
SRU replacement	10 minutes
	20 minutes maximum

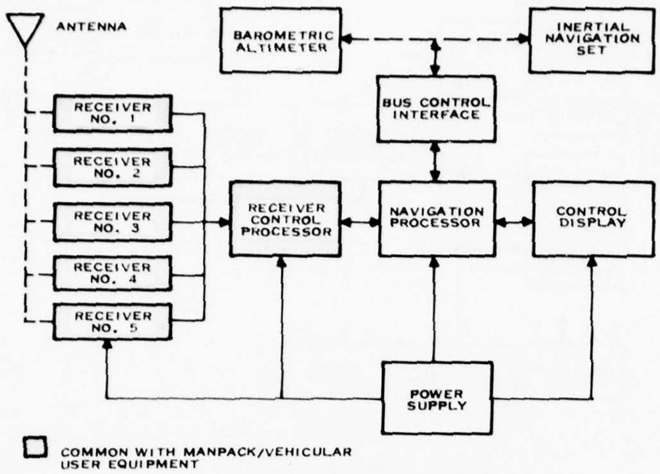


Figure 12. High Dynamic User Equipment Block Diagram

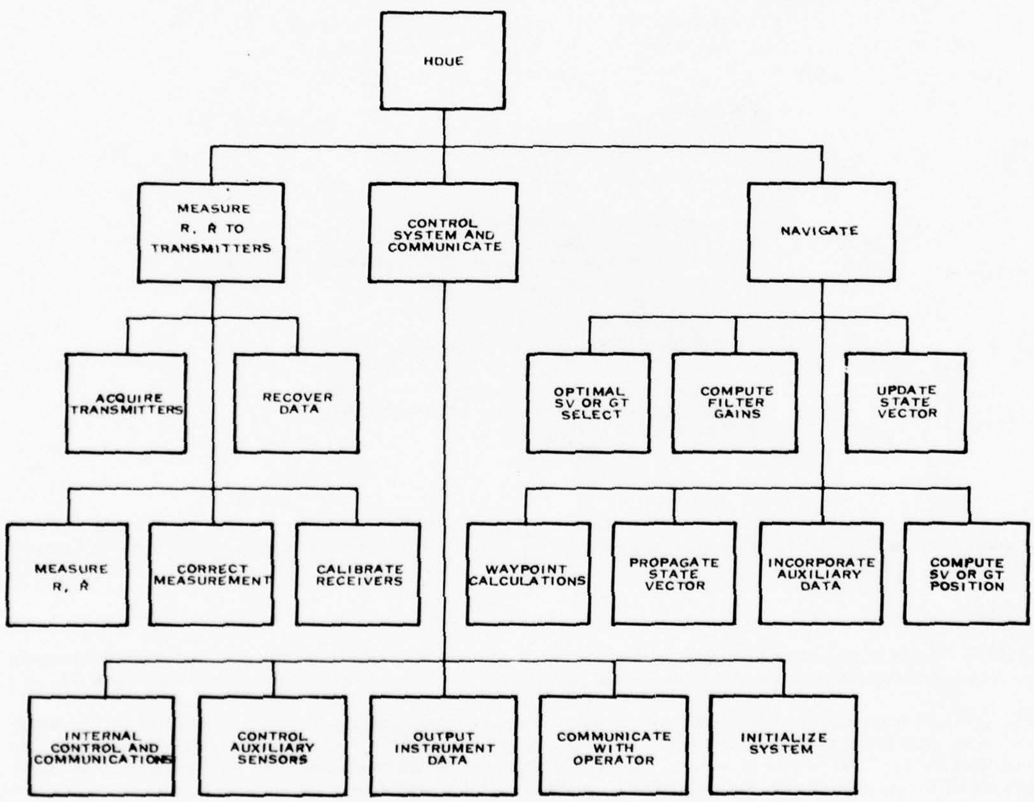


Figure 13. HDUE Major Modes of Operation

The HDUE set tracks four satellites continuously using four receiver channels. The fifth receiver channel is used for SV acquisition, L_1/L_2 ionospheric correction measurement, and reading of data. The HDUE tracks both L_1 and L_2 satellite frequencies and uses both C/A and P codes during the acquisition process. Navigation calculations are made using an 11-state Kalman filter.

The HDUE set can handle two antenna inputs (inverted range or satellite antenna) and switch any of these two inputs to any one of the five receiver channels (2 × 5 matrix switch). Input initialization data (position and time) are entered by means of the

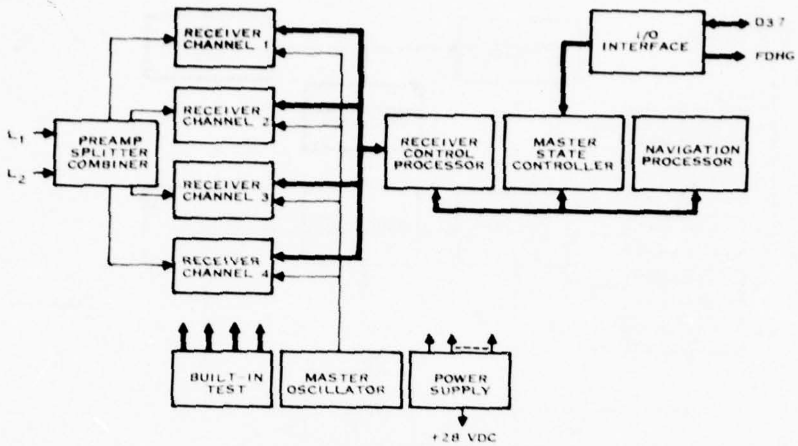


Figure 14. MBRS Block Diagram

TABLE 11. MBRS OPERATIONAL CHARACTERISTICS (POSITION, VELOCITY)

Mission Sequence

Prelaunch

- $C/N_1 = 33$ dB-Hz
- $T_{ACQ} = 120$ seconds

Launch

- $C/N_1 = 33$ dB-Hz
- $T_{ACQ} = 10$ seconds
- Probability of acquisition = 0.95

Powered flight

- $\dot{R}_{MAX} = 25{,}000$ feet/second
- $\ddot{R}_{MAX} = 10\,g$
- $\dddot{R}_{MAX} = 9$ meters/second^3 ($T_{MAX} > 0.25$ second)
  - $\sigma R = 4.4$ feet ($C/N_1 = 29$ dB-Hz)
  - $\sigma \dot{R} = 0.012$ meter/second (1 second average) ($C/N_1 = 29$ dB-Hz)
  - $\sigma R = 8$ feet ($C/N_1 = 25$ dB-Hz)
- $T_{ACQ} = 10$ seconds ($C/N_1 = 33$ dB-Hz)
- Probability of acquisition = 0.95
- $J_{STAGE} = 50\,g$/second, $100\,g$/second, and $300\,g$/second ($T < 0.1$ second)

Post boost

- $\sigma R = 4.4$ feet (C/N_1 29 dB-Hz)
- $\sigma \dot{R} = 0.012$ meter/second (1 second average) ($C/N_1 = 29$ dB-Hz)
- $\sigma R = 8$ feet ($C/N_1 = 25$ dB-Hz)
- $T_{ACQ} = 10$ seconds ($C/N_1 = 33$ dB-Hz)

keyboard on the CDU. Output data (position, velocity, time, etc.) are displayed on the CDU. The same receiver and processor common modules are used in the HDUE as are used in the MVUE. Software is programmed in both Fortran and 990 assembly language.

VIII. MBRS

The MBRS consists of one major LRU that contains the receiver, receiver controller, data processors, and power conditioning modules, forming a four-channel, continuous-tracking receiver configuration. A block diagram of this set is shown in Figure 14.

The major characteristics of the MBRS are shown in Table 11 and the operational sequence is shown in Table 12. The MBRS interfaces directly with the digital control unit (DCU) computer on the Minuteman missile. All initial inputs of position and time are entered from the DCU. All outputs of position, velocity, and time from the MBRS are sent to the DCU or telemetry assembly. The MBRS is used to provide navigation instrumentation for postflight test data evaluation of the missile trajectory.

The MBRS has improved features in the receiver area. These include a digital oscillator to track the increased missile doppler frequency and provide increased position and velocity accuracy. Also, the carrier tracking loop circuitry is implemented in a special digital processor.

MBRS tracks both L_1 and L_2 satellite frequencies and uses both C/A and P codes during the acquisition process. Navigation calculations are made using an 11-state Kalman filter.

The antenna inputs at both L_1 and L_2 (1575.42 MHz and 1227.6 MHz) frequencies. The missile 28-Vdc prime power is regulated and conditioned for use in the MBRS.

All three GPS Phase I user sets use common receiver/processor modules to the maximum extent possible. The extent of module commonality between sets is shown by the module comparisons listed in Table 13. Unique modules were designed to cover the individual differences and special requirements of each GPS set. These unique modules are listed in Table 14. The greater use of common modules will increase the total production volume for each module and, hence, drive module cost down. Maximum use of standard replaceable modules will continue in the next generation of GPS user equipment.

TABLE 12. MBRS FUNCTIONAL OPERATIONS

Turnon	Track GPS signals
Self-test	Read SV data
Receive initialization data	Establish position
Send initialization data for verification	Send R, Ṙ, T to telemetry
Acquire selected GPS signals	Perform L_1/L_2 switching sequence

TABLE 13. PHASE I GPS–COMMON MODULE COMPARISONS

Module Name	Number of Units/Modules/Cards		
	HDUE	MBRS	MVUE
Wideband	5	4	1
Narrowband	10	12	2
Output	5	–	1
Frequency synthesizer	5	–	1
Code generator	5	4	1
Clock	1	1	–
MPM	3	4	1
4K DMM	17	4	1
16K PMM (MBRS legacy)	–	4	2
IBIM	2	4	–
FPAU	2	2	–
FPAU	1	1	–
FPAU	1	1	–
FPAU	1	1	–
FPAU	1	1	–
FPAU	1	1	–
CRIM	3	4	1
Receiver test	1	1	1
Distribution	1	1	–

TABLE 14. PHASE I GPS–UNIQUE MODULE COMPARISONS

Module Name	Number of Units/Modules/Cards		
	HDUE	MBRS	MVUE
Antenna	2 (GFE)	–	1
Filter/preamplifiers	2 (GFE)	1	1
Antenna switch	1	–	–
CDU	1	–	1
Master oscillator	1	1	1
Power supply (regulated)	4	1	1
EIOM	–	–	1
TACAN/clock I/O	–	–	–
IFM	–	1	–
SCM	–	1	–
Frequency synthesizer	–	4	–
Output module	–	4	–
DRIM	1	–	–
SBIM No. 1	2	–	–
SBIM No. 2	2	–	–
DBEM	1	–	–

IX. REFERENCES

1. Damoulakis, J.N., and Cocharo, G., "A Double-Stage Piecewise Recursive Filter for Real-Time Applications" (pending SAMSO approval).

2. Damoulakis, J.N., and Hinderer, J., "A GPS-HDUE Configuration With Increased Processing Power on Range and Range-Rate Measurements," Position Location and Navigation Symposium, Institute of Electrical and Electronics Engineers (IEEEO, 1–3 November 1976, San Diego, California).

3. Upadhyay, T.N., and Damoulakis, J.N., "A Real-Time Suboptimal Filtering Algorithm for GPS/MVUE Navigation Application" (in preparation).

GPS RECEIVER OPERATION

by

Burton G. Glazer
Division Vice President and Director
Communication Systems Operation
Advanced Products Division
Magnavox Government & Industrial Electronics Company
2829 Maricopa Street
Torrance, California 90503

ABSTRACT

Different receiver configurations are suitable for applications having differing levels of dynamics of the Host Vehicle and interference environments. All configurations must be capable of accomplishing certain fundamental operations: Satellite selection, signal acquisition, tracking and measurement and data recovery. After correction for propagation effects, the signal time-of-arrival measurements are used to obtain the navigation solution. Limits of performance of receivers are described.

I. ARCHITECTURE

A GPS User Equipment is comprised of four principal components: Antenna, Receiver, Computer and Input/Output devices.

The antenna in most cases is a relatively simple element providing approximately isotropic gain from the zenith to the horizon at one or both of the GPS frequencies. Since the signals are circularly polarized, a conical spiral or variation thereof is suitable. Where high performance is required, particularly near sources of interference, a more elaborate antenna system such as a steered-beam phased array or null steering adaptive array may be used. Antenna placement on the Host Vehicle is important in two regards. There should be a clear view of the whole sky; shadowing of some of the satellites can result in degraded performance. A potentially significant error source is multipath, particularly of the stationary sort produced by reflections from surfaces near the antenna. In the extreme, measurement errors of tens of meters on the P signal and hundreds of meters on the C/A signal can arise from such reflections. Both of these considerations favor antenna placement at the highest point on the vehicle. In the majority of vehicles it is impractical to locate the receiver there, and so a preamplifier must be used to drive the cable to the receiver and avoid losses due to cable attenuation and interference pick-up.

The primary decision in the selection of receiver architecture for each application is the number of signals to be processed simultaneously by the receiver. Each satellite transmits three signals (C/A and P on the L1 frequency and P on the L2 frequency) and in the fully deployed system there will be as many as eleven satellites in view at a time. While a thirty-three channel receiver capable of tracking and measuring all of these signals simultaneously might be useful for a monitor station, users can obtain nearly optimum performance with one to five channels. The minimum receiver provides only a single channel capable of recovering one C/A signal at a time. More elaborate receivers process five signals at once, where each may be chosen to be a C/A or a P signal on the L1 frequency or a P signal on the L2 frequency. The number of channels and elaborateness of the receiver structure is primarily dependent upon the maneuverability of the Host Vehicles and secondarily upon the accuracy and interference resistance required.

A computer is a necessary part of a GPS User Equipment. In ground based radio navigation systems (LORAN, VORTAC, OMEGA) after making the signal measurements, the navigation computation may be completed manually using charts overprinted with lines of position corresponding to measured values of phase or time difference. This is not the case for satellite-based navigation systems: The charts would have to be reprinted every millisecond. Sets thus far publicly described have used one or a few 16-bit minicomputers.

Of interest is the division of functions between the receiver and the computer. Such functions as carrier tracking loop filtering and data bit detection have been accomplished sometimes in the receiver, sometimes in the computer. It has been seriously suggested that the signal be sampled and quantized prior to correlation and that all subsequent processing be accomplished in a high speed computer. Computer functions always include control of the receiver, selection of satellites and signals to be utilized, correction of measurements for propagation effects, computation of position and velocity in the desired coordinate system and communication with other systems and the operator through the Input/Output devices. Typically, programs for the GPS user computer are of the order of 30,000 words in length.

Input/Output devices are as widely varied as the applications of GPS. When operated by a human operator, compact keyboards and alphanumeric displays provide the interface. GPS sets can drive cockpit instruments or generate data link messages for position reporting by radio. The GPS set does require certain information to facilitate start-up: Approximate satellite and Host Vehicle locations and time. Various methods of introducing this data such as keyboards, cassettes, radio links or data busses may be used. If antenna beam steering is to be accomplished, or the location of a point on the vehicle other than the GPS antenna is desired, then attitude information must be fed to the set. Although the GPS set produces elevation above a geodetic reference as an output under normal operating conditions, it is considered useful to provide barometric altitude as an input to assist in obtaining the first fix and to provide extra information should satellites become unavailable due to shadowing or other temporary outages.

The "Z-Set", shown in Figure 1, provides a current example of a nearly minimal GPS User Receiver suitable for non-combat military aircraft. Unit manufacturing cost in small-scale production is about $15,000. This set utilizes only the C/A signals on the L1 frequency to provide position accuracies superior to all existing medium and long-range radio or inertial navigation systems.

The major modules of this set are indicated in Figure 2. The antenna/preamplifier unit is best located either in a tail cap or on the top of the fuselage, just aft of the cockpit. The preamplifier incorporates pre-selection filters and about 30 dB of gain at the L1 frequency, 1575 MHz. The overall noise figure is about 4 dB. The receiver/processor unit may be located in any convenient place. The reference oscillator is a good quality crystal oscillator in an oven, with particular care to minimize sensitivity to vibration since phase noise and short-term drift can adversely affect performance. From this oscillator are synthesized the several local oscillator frequencies required by the receiver and the basic time pulses to be counted by the User Time Clock (UTC) module. This count is the time reference against which the signal arrival times are noted. In the RF-IF module, the signal is further filtered, amplified and translated down in frequency. In part, the translation is determined by a voltage controlled oscillator which is driven by tracking loops in the baseband module to offset the doppler shift. This VCO frequency, suitably scaled for the ratio of the C/A code rate to the L-band carrier frequency, also contributes to the synthesis of the clock which drives the C/A coder. This is done in such a way that if the receiver is phase or frequency locked to the L-band carrier, the C/A coder clock will be correct in frequency, requiring only a correct initial phasing to the incoming signal to remain aligned with it thereafter. The baseband module contains the detector circuits for carrier frequency, phase and C/A code error sensing as well as power detection circuitry for recognizing initial alignment and a data demodulator. Upon the occurrence of certain events in the C/A code generator, the UTC time count is strobed and transferred to the computer via an I/O module. The I/O modules contain various buffers, drivers and handshaking logic needed to process the digital signals into and out of the computer. The computer is comprised of four modules: A CPU and three memory modules. This CPU is designed around the LSI-11 central processor, but is augmented beyond that well known computer by the addition of microprogrammed instructions to facilitate high precision floating point computations. Through an interface module, this set can accept a digital altimeter reading, drive cockpit instruments or provide information to other navigation equipments.

II. STARTING OPERATIONS

The first operating function is the selection of the satellites to be used in the navigation solution. Fundamentally, the set is to determine the values of four unknowns: Three position coordinates and time. Accordingly, four measurements will be needed, usually the time-of-arrival of signals from four different satellites. In some situations, as when altitude or time is very accurately known, fewer measurements will suffice.

The set must be provided with information regarding the location of each of the satellites as a function of time. For the purposes of selection and signal acquisition, this "almanac" need be accurate only to a few kilometers, and it is estimated that almanacs will be usable for a week or more, so that a set which is regularly used can retain the almanac in a nonvolatile memory from one usage to the next. Each satellite transmits the current almanac as part of the navigation data message, allowing users to update their stored almanac. For a truly cold start in which the set does not contain a valid almanac, two approaches are possible. The almanac may be transferred from an active GPS receiver via data link, cassette or the manual keyboard. Alternatively, a "search-the-sky" approach may be used in which the set simply tries to acquire the C/A signal of each satellite in turn, without prior knowledge of satellite visibility or doppler. Unless the receiver includes a matched filter for expediting the synchronization process, the search-the-sky process can take many minutes to complete.

When provided with a valid almanac, an approximate knowledge of its own position and time-of-day, the computer can execute a satellite selection algorithm. To minimize the sensitivity of the position solution to measurement errors, the satellites as viewed from the user should have the largest possible angular separations. After excluding those satellites which are or soon will be, below the horizon, subsets of four can be tested for angular separation. Other criteria may also be introduced in the selection process such as satellite signal quality or avoidance of the use of low elevation satellites whose observation is most subject to propagation error sources. Subsequent to the selection of the initial subset of satellites, or "constellation", the selection should be reviewed every few minutes and revised when necessary to maintain minimum navigation error as the geometry of the constellation changes.

The next question is the acquisition of the signals from the selected satellites. The normal method of signal acquisition is to synchronize to the C/A signal and then, when necessary, transfer to the P. The computer must designate to the receiver not only the satellite to be acquired, but also an estimate of the expected doppler shift on the signal. This is useful because the doppler range of ± 5 KHz (for a slow moving user) is so large that the signal-to-noise ratio in the corresponding bandwidth is less than unity. Under this condition, it is more effective to subdivide the frequency uncertainty and search sequentially, thus an estimate of doppler can significantly reduce search time. The usual method is to correlate the incoming signal against a local replica consisting of the chosen C/A sequence modulated on the receiver local oscillator. The time phasing of the C/A sequence is varied slowly until the post-correlation power exhibits a rise above that which might be attributable to noise alone.

Having established synchronization of the pseudonoise sequences, tracking is begun in both code sequence timing and carrier phase. Once the tracking loops pull in, the data format features (bit edges, word starts, subframe starts and Z counts) may be recognized to provide unambiguous time-of-arrival and time-of-day indication. While the almanac is sufficiently accurate for acquisition, much more accurate information on satellite position and the offsets of its clock is needed to achieve the desired navigation accuracy. These data, called the ephemeris, are contained in about twenty 24-bit words which are part of the data format transmitted by each satellite. Having recognized the format identifiers, the ephemeris can be recovered. Although the 6 parity checks accompanying each of the 24-bit words provide sufficient redundancy for error correction, it is advisable to use the redundancy for error detection only, thereby

obtaining extremely high confidence that erroneous ephemeris will not be accepted into the data base for the navigation solution.

III. MEASUREMENTS

The primary measurements made by a GPS receiver are the times of arrival of the satellite generated signals. Arrival time is measured with respect to the receiver clock which is a stable oscillator and counter. The time and frequency offsets of this clock from GPS system time and true frequency are not critical, since these will be determined from the navigation solution. Of considerable importance is the stability of the oscillator, particularly in receivers which observe the signals sequentially.

Signal-to-noise considerations preclude direct measurement of the arrival time of a particular PN code element edge. A replica PN sequence generator is caused to run in phase with the modulation of the incoming signal by a tracking loop. Simultaneously or alternately, the powers recovered by correlation of the incoming signal against slightly advanced and delayed versions of the local replica are compared to provide the tracking error indication. The arrival time may then be observed by comparing the replica code timing with the receiver clock timing. Either the clock time can be observed upon the occurrence of certain events of the replica (such as the first PN edge of a data word) or the phase of the replica relative to a hypothetical replica driven by the receiver clock can be observed at certain events of the receiver clock, such as 1 second ticks. While substantially equivalent, there are significant hardware and computational differences between the two schemes. The first technique appears to be slightly more economical for single channel receivers, the second technique is preferrable for multi-channel or aided receivers.

These measurements, loosely called pseudoranges, must be adjusted for propagation effects. The dominant departure from a simple free space propagation model is the additional signal group delay due to the passage of the signal through the ionosphere. At the present state of knowledge, simple prediction models of the ionosphere do not yield very accurate results: ten or twenty meters error under adverse conditions. There is some hope that improved knowledge, obtained in the next few years from GPS itself, will give rise to an improved prediction technique. The additional delay in traversing the ionosphere is best estimated by observing the difference in arrival time of the P signals on the L1 and L2 frequencies. Since along a given path through the ionosphere at a given time the group delay varies inversely as the square of the carrier frequency, the two frequency observations suffice to determine the proportionality factor. The delay changes rather slowly, at most about a nanosecond per minute, so that only occasional dual frequency observations are required. Because two measurements are needed to compute the delay correction and the two frequencies are close, there is a magnification of measurement noise in obtaining the correction, and so some filtering of this measurement is desirable. Tropospheric delay also requires correction. The use of a simple altitude and elevation angle dependent model is sufficient to reduce the error to negligible levels. While there are some techniques for avoidance of multipath effects and for recognition that a reflected signal rather than the direct signal is being tracked, in those cases when the multipath signal is within 1 or 1-1/2 PN chips of the direct signal there is no practical way to correct the resulting erroneous measurement.

The satellite clock error (from system or universal time) is monitored and modeled by the Control Segment. Coefficients for a correction computation are sent to each satellite for subsequent retransmission to the users. Thus, the raw pseudorange measurements are compensated for ionosphere and troposphere effects and then for satellite clock offsets prior to beginning the position computation.

In some sets incremental as well as whole values of pseudorange are measured. When signal-to-noise ratios permit phase locked tracking of the received carriers, it is possible to count beats of the difference frequency between each received carrier and a hypothetical carrier synthesized from the receiver's frequency standard. While this sort of observation is sometimes called a doppler measurement, it is really a differential pseudorange measurement in which the change in pseudorange from the start to the end of the counting interval is observable with accuracy and resolution of a fraction of a carrier wavelength. These measurements are useful when a fast velocity measurement is wanted, or in integrated GPS inertial systems or in single channel sets for dead reckoning between pseudorange observations of each signal.

IV. POSITION COMPUTATION

The position of the satellite is computed from the ephemeris parameters received from the satellite. The parameters appear to describe a Keplerian ellipse with some correction terms for oblateness of the earth and rotation of the orbit plane. To have a complete analytical expression for a whole orbit to the desired accuracy (better than 1 meter) would take far more parameters and impose a huge computational load on the users. The transmitted parameters, precomputed by the Control Segment, are actually a best fit of the corrected ellipse to the true orbit which meets the desired accuracy over one hour. When a satellite is first acquired, this data is collected. In the event that that satellite remains in the constellation for more than sixty or ninety minutes, a new ephemeris should be collected. At each time of pseudorange measurement, the corresponding position of the satellite is computed by inserting the value of time in the equations of the corrected ellipse and converting to a more convenient coordinate system. For most purposes, we prefer a Cartesian earth centered, earth fixed coordinate frame for performing the basic position solution.

At this point, in the simplest case, we have four unknowns (three coordinates of user position and user clock time offset) and four equations each involving the measured, compensated pseudorange (the scalar distance of user to the satellite plus the clock bias) all units adjusted to meters. The equations are non-linear, but capable of solution by any of a number of techniques suitable for computer execution.

Usually, and always to some advantage, a more sophisticated viewpoint is taken which allows for utilization of additional (or fewer) inputs and thereby improving the solution in terms of accuracy or reliability under unfavorable conditions. Even in the absence of external inputs there is a benefit to be obtained if the GPS user clock is stable. If, after a few of the four equation solutions it become possible to predict what the time offset of the user clock will be, the number of unknowns is reduced to three.

If four measurements are still available, the redundant measurement can be used to improve the accuracy of the solution. Alternatively, the fourth measurement could be sacrificed to allow the associated hardware (in a multichannel receiver) or time (in a sequential receiver) to be used for other purposes such as dual frequency measurements or acquiring a new satellite when the constellation is to be revised. The loss of the fourth measurement may be inadvertent, as when the wingtip of a banking aircraft is shadowing a satellite.

The usual formulation is to make the position solution part of a Kalman filtering operation. The user state vector might usefully contain nine terms (give or take a few) such as the three position coordinates, the three components of velocity, clock time offset, clock frequency offset and altitude. The observation vector might include pseudoranges, differential pseudoranges, altimeter and other air data, accelerometer outputs, or position indications from other navigation systems.

V. NAVIGATION COMPUTATION AND OUTPUTS

User position in Cartesian earth centered coordinates is rarely the acceptable end product. Each application has its own preferences. Latitude, longitude, and MSL altitude for position is commonly desired, but sometimes easting and northing in the military grid reference system is demanded. For navigation, the preference is often for steering instructions to a specified destination in terms of range and bearing or crosstrack error and time to go or other forms. All of these involve additional computation and program, parameter and data storage but contribute greatly to the utility of the set.

The form of outputs is also application dependent. Alphanumeric displays in incandescent, LED or LCD form is often adequate. In survey work, hard copy is desired. In airborne use, the choice may be for moving needle indicators. For aircraft and ships, automatic position reporting via digital data link is a likely future requirement.

VI. ACCURACY

In GPS there are many error sources, some of which are difficult to characterize and simple one-figure accuracy statements must be carefully interpreted.

From the Space and Control Segments come satellite position uncertainties and satellite clock drift. These are projected to be 1.5 meters and 1.0 meter (equivalent) respectively, in their contribution to pseudorange error.

If the ionosphere effect is predicted with current knowledge, the error is very much dependent on time of day, solar activity, geomagnetic latitude and other factors determining the condition of the ionosphere. On a bad day, the errors may be 30 meters, although on an overall average it may be only 3 meters. The dual frequency measurement method will give 1 to 3 meters depending upon signal conditions and receiver design.

The residual after tropospheric correction will be 1 meter.

Characterization of multipath errors is virtually impossible, since it depends upon specific antenna locations and the location and surface conditions of nearby objects. It is possible to create situations with multipath induced errors of tens of meters for P code (hundreds of meters for C/A code), but with reasonable care and luck, 1 meter or less will be far more typical.

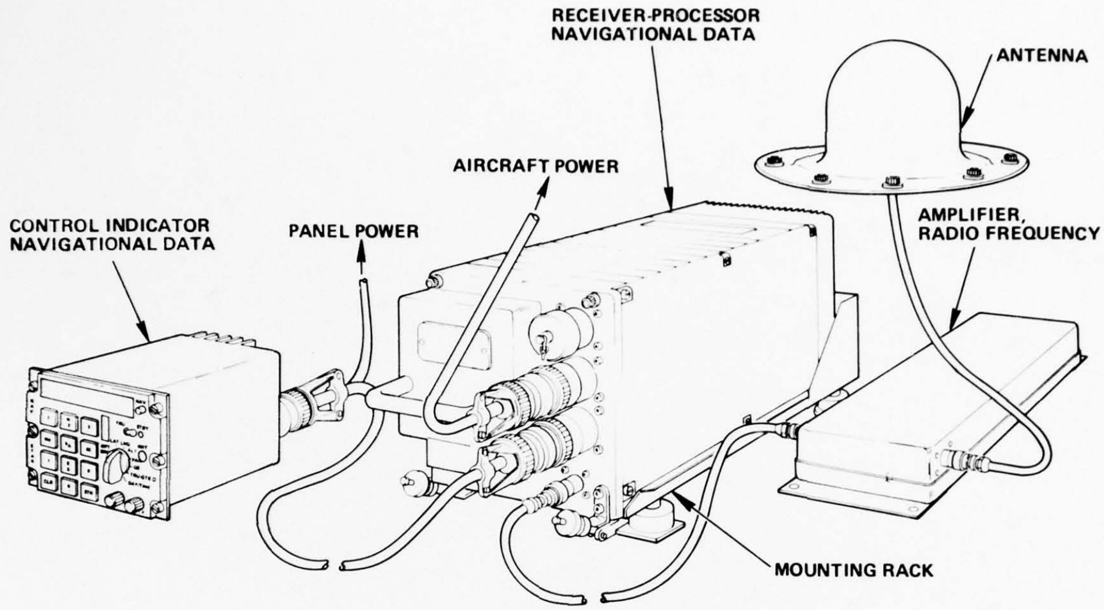
Receiver measurement error is very much a function of the level of interference at the antenna and to some degree upon the Host Vehicle dynamics, both the actual and the design values. Under benign conditions the measurement error will be under 1 meter for P code (10 meters for C/A) and twice that as the noise and dynamic limits are approached, beyond which point no measurements are available at all.

Receiver mechanization errors arising from such sources as offsets in measurement circuits, quantization noise, computational approximations and truncation errors can be held to 1 meter with careful design.

Without being precise about operating conditions or confidence levels, these errors add (in the square root of the sum of the squares) to 4 to 6 meters in pseudorange for the case of P code, dual frequency sets. In solving for position from the pseudoranges these errors are multiplied by a factor of 1.5 to 2 for most times and places, assuming no substantial blocking of the antenna. A projection of "10 meters rms or better, almost always" is supported by early test results utilizing the first experimental GPS satellite.

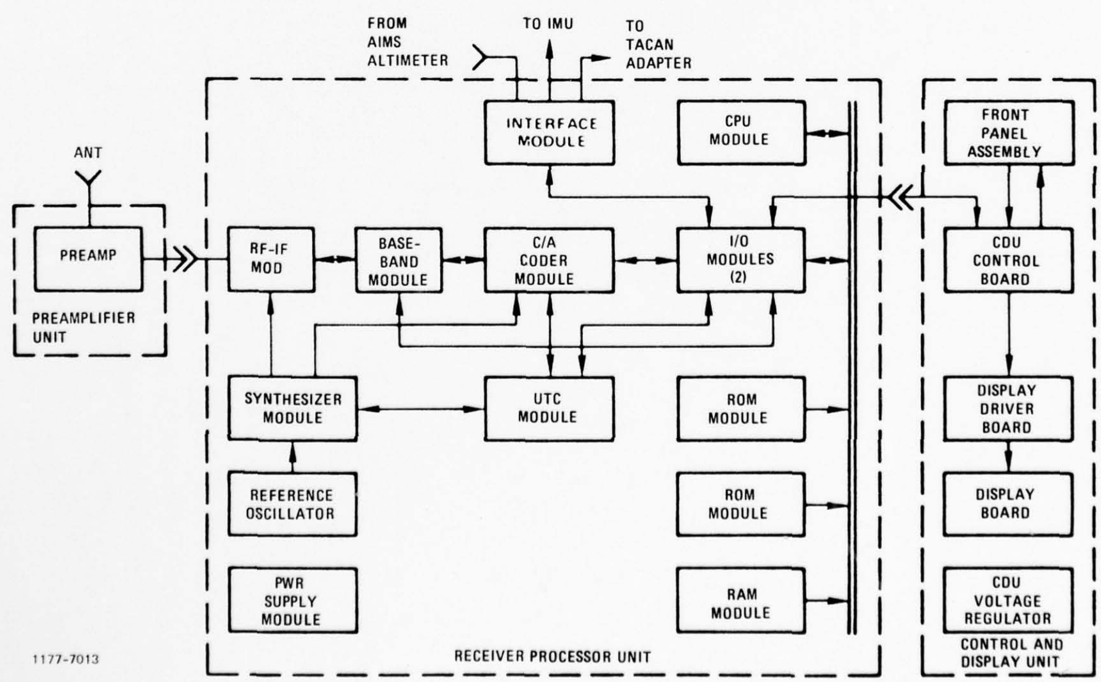
For a minimum cost set such as the Z which uses only C/A code on one frequency, the dominant error source is the ionosphere when it is bad, and receiver measurement noise when it is good. Position accuracy will be about 30 meters under most conditions, but double that under extreme ionosphere situations.

The error analyst has often been asked whether there is benefit to be derived from using more than the minimum three or four satellites. Certainly more measurements are helpful, but under the practical constraint (due to receiver hardware or computational limitations) of allowing only a given number of measurements per minute, the question is more interesting. If all of the error sources were unbiased and independent of the choice of satellite then the best approach would be to select the minimum number of satellites which are geometrically well spaced and whose signal strengths are good and then to use all of the measurement capacity on them. If, on the other extreme, the errors were dominated by sources associated with particular satellites, such as satellite ephemeris errors or multipath reflections which were strong in certain directions, then the best strategy would be to gather measurements on most of the satellites in view. At this time the former situation is expected to be the more likely, but this opinion is subject to revision as the testing proceeds during the coming year.



1177-7101

Figure 1. Set Z Navigational Set



1177-7013

Figure 2. Set Z Major Modules

PHASE II GPS RECEIVER DESIGN PHILOSOPHY

Frank E. Butterfield
Robert E. Sung
The Aerospace Corporation
Post Office Box 92957
Los Angeles, California 90009

SUMMARY

Phase II User Equipment developments are the second step of a three phase program for introducing this navigation system into operating elements of the Military forces.

Phase I was largely restricted to Advanced Development Models for proof of system validity. There are two exceptions: a Generalized Development Model of a High Performance User Equipment and a prototype of a low cost model. Phase III will be production from prototypes resulting from Phase II.

Design to Life Cycle Cost is the process being used to guarantee high value apparatus.

For greatest utility the Phase II designs, the prototypes for Phase III production, are being engineered for intimate integration into a wide variety of host vehicles.

1.0 INTRODUCTION

This chapter discusses the development of the Global Positioning System (GPS) User System Segment. First, the development objectives, program plan and a brief review of development activities will be presented. This is followed by a discussion of user requirements, host vehicle interfaces and integration, definition of user equipment set configurations and the elements of a Design to Cost (DTC)/ Life Cycle Cost (LCC) program. The GPS Phase II User Equipment (UE) DTC/LCC program objectives and tasks required to accomplish these objectives are described brieflv. Finally, the modular design concept,which embraces the use of a limited inventory of standard function replaceable hardware and computer program or software modules to implement specific operational building blocks to be uniquely arranged for the synthesis of differing UE sets,is presented along with the flexible interface module concept and the "form, fit and function" approach to integration.

1.1 Purpose

Most of the design concepts presented have evolved from the collective development efforts on the parts of the government and industry. These concepts will be further studied, analyzed and probably modified during the Phase II UE development process in order to yield a design which minimizes life cycle cost, while providing positioning and navigation services at acceptable performance levels. It is hoped that through an overview of the objectives, requirements, constraints and design approaches, this chapter contributes to an understanding of the GPS Phase II UE design and development philosophy. The Phase III issues and options which pertain to the production and support of the GPS user equipments will not be discussed here at any length although they will be identified at the onset of Phase II effort and continuously addressed through the acquisition cycle.

2.0 OBJECTIVES AND PROGRAM PLAN

2.1 Phase I

The primary objective of GPS UE development is to support and implement the overall acquisition approach selected for GPS, which is a three-phase, evolutionary, design-to-cost development and test program leading in successive phases to a world-wide, operational capability providing extremely accurate three-dimensional position and velocity information and system time to suitably equipped users anywhere on or near the earth. The three phases of the GPS development allow the system to evolve with each phase building and expanding on the previous phase in an integrated and cohesive manner. The program initiation decision by the Defense Systems Acquisition Review Council was to proceed with Phase I, Concept Validation, to confirm the basic system concept, demonstrate the capabilities of the preferred GPS design, and provide in-depth information on system cost and military value of such a system. In addition, the necessary answers to questions of design, performance, schedule, life cycle cost and applications are to be obtained prior to the DSARC II ratification decision.

The UE development objective during Phase I is oriented towards the support of Development Test and Evaluation (DT&E) of Advanced Development Models (ADMs) of user equipment to validate the GPS concept and to gather performance, engineering and cost data of various receiver design approaches. Towards this end, a total of six different user equipment models have been designed and fabricated. These are Sets X, Y, Z and Manpack built by Magnavox Advanced Products Division and High Dynamics User Equipment (HDUE), Manpack/Vehicular User Equipment (MVUE) by Texas Instruments. In addition, one Generalized Development Model has been developed by Collins Avionics Division of Rockwell International. It will be tested to verify jam resistant design alternatives. This technology effort, conducted by the Air Force Avionics Laboratory (AFAL), will contribute to the identification of preferred user equipment design for Phase II.

The major emphasis in Phase I is on the receiver design. There is a range of receiver opportunities relative to number of channels, techniques of signal detection and tracking, techniques of construction, and environmental restrictions. One of the Phase I objectives is to encourage experimentation for a limited time and continue to keep these technology options open. Since Phase I will encompass

the first of three design-build-test-design cycles to determine preferred user equipment configurations at an affordable cost, another objective is to ensure legacy in the development hardware and software. Engineering activities have been oriented towards design to cost management and life cycle cost analysis at the design engineering level. The development plan reflects the iterative "design-test-improve" concept in Design to Cost engineering.

2.2 Phase II

A decision to continue into Phase II Full Scale Development could be made at DSARC II in early 1979. The Phase II efforts will complete the Initial Operational Test and Evaluation (IOT&E) of prototypes of production user equipment for a broad spectrum of Army, Navy, Air Force, Marine Corps, Defense Mapping Agency and other users. The emphasis now shifts to developing a family of user equipment sets for minimum life cycle cost which will meet the users' performance requirements while satisfying weight, power, size and other interface constraints on the integration of the GPS equipments into their host vehicles.

The procurement for Phase II of GPS navigation equipment is divided into two stages - Phase IIA and Phase IIB. Phase IIA due to start in June 1978 will include a series of trade studies and analyses by four competitively selected contractors to identify the preferred GPS navigation system architectures and support concept. Minimization of the life cycle cost of the total inventory of multi-service user equipment is the major objective of these studies and analyses. Phase IIB, commencing after DSARC II, will have two contractors selected from among the Phase IIA contractors to continue with design refinement, fabrication of prototype equipment and extensive testing.

2.3 Phase III

Phase III, Production/Development, will develop the full global capability of the system at a rate to meet a planned Initial Operational Capability date in 1984. The UE development for Phase III will be for the production of operational GPS user equipments. The initial Phase III contractor will be selected from the competing Phase IIB contractors. IOT&E results, life cycle cost, commitment to support cost guarantee and reliability improvement warranty (RIW), producibility and management capability will be given heavy consideration. A leader/follower (s) procurement concept is envisioned for the production of the preferred Phase IIB design with the leader qualifying a second production source (s) within 18 months after the first production contract award. The leader will retain total system performance responsibility to include responsibility for any system level or module level support cost control.

3.0 USER REQUIREMENTS

3.1 System Requirements

Since the early 60's, a great number of military missions and applications were analyzed in defining the objectives for the GPS. These applications range from precision weapon delivery through search and rescue to geodetic survey. Most of these applications fall into the following categories: weapons deliveries, enroute and terminal navigation, range instrumentation, targeting and survey, and all weather or constrained maneuver operations for the transportation of men and materials.

From these mission requirements are extracted the desired system characteristics for GPS which are summarized in Table I. As these capabilities can be directly translated into military gains in terms of increased weapon system effectiveness, faster response, greater flexibility in force deployment, and enhanced probability of mission accomplishment, new applications and potential users continue to be identified.

Table I. GPS Characteristics

1.	World-wide coverage (continuous)	8.	Not line of sight limited
2.	Accuracy compatible with mission of user	9.	Denied to unauthorized use
3.	All weather (and not daylight dependent)	10.	Jam resistant
4.	Real time (and rapid acquisition)	11.	Cost effective
5.	Unlimited number of users	12.	No altitude dependence
6.	No user radiation	13.	3 dimensions (position and velocity) and time
7.	Common grid	14.	Users independent

3.2 Diversity of User Requirements

The GPS user community is extremely diverse, representing many different host vehicles and many different missions. Each user requires certain specific functional capabilities and certain minimum acceptable performance levels, but attaches different weights to various specification items and penalties to his specific host vehicle. It would, of course, be hopelessly expensive and logistically impossible to tailor-make an equipment for each and every user. Hence there has been a concerted effort to define several user classes (groups) by the order of importance placed on requirements and the tolerable range of performance for each specification item in each class.

3.3 Equipment Classification

It has been noted earlier that although a wide variety of applications require differing performance levels, some of these differences can be accommodated economically by different configurations of elements which themselves satisfy the requirements of several configurations. A good example is the difference between equipments which track four or more signals simultaneously and equipments which

track signals in turn. While the first employs four or more signal processing, code generating, signal tracking and range measuring circuit groups, the second uses only one at a time and indeed can use most or all of a single group for all its signals. Another example is the difference between equipments intended to support users with significant dynamics (aircraft) and those intended to support users with low dynamics (survey equipment). Similar parts are appropriate with bandwidth differences produced under computer program control.

These differences have lead to the identification of classes of equipment, generally performing to different specification levels, built from identical circuit elements or "modules". It is important in obtaining low life cycle costs to restrict the variety of parts in the overall system and to manufacture as many of identical design as possible. However, one obstacle to standardization of modules seems to be the different operational test requirements of the military services which use GPS*.

While the operational test requirements may differ between military services, the classes of equipment or performance requirements are generally not segregated by services. All services fly rotary wing aircraft. All services have transport tasks which generally involve similar dynamics, accuracy requirements and volume restrictions. Consequently classes usually do not correspond to service peculiar practices or even completely consistent environmental restrictions.

3.3.1 Early Classification

The early categorization was as follows:

Description	Class Designation
Severe Dynamics in Jamming	A
Severe Dynamics W/O Jamming	B
Low Dynamics, No Jamming; Low Cost	C
Wheeled and Tracked Vehicles	D
Manpack	E
Ships (all types)	F

This division was too fine for the Phase I program and was replaced with

Description	Class Designation
Full Dynamic Performance	X
Moderate Dynamic Performance	Y
Low Cost	Z
Manpack/Vehicular	Manpack

Some months into the program it became apparent that X and Y had two embodiments - operated with inertial navigators and stand alone. Therefore the term "X-aided" entered the vocabulary. Further into the program, new items were added, including the HDUE and MVUE,GPSPAC (a spaceborne set), and the M-set for missile applications.

3.3.2 Phase II Classification

In Phase II, where prototypes for production are required, classes are oriented around common applications and host vehicles of common performance capability as shown in Table II.

Table II. Installation Categories

Host Vehicle Category		Typical Vehicles
1.	Manpack/Vehicular	Man, Jeep, Tank
2.	Helicopter/Recon	UH-I, RH 53D, OV-ID, AH-1S
3.	Fighter/Attack	A-4, A-6, F14, F15, F-4
4.	Transport/Tanker/ASW	C-141, C-5A, E-3A, P-3C
5.	Strategic Aircraft	B-52
6.	Surface Ship	CV, MSO
7.	Submarine	Submarines
8.	Trainer/Transport	T38, C135, F-5E
9.	Austere Navigator	To be identified.

3.4 Phase II Requirements

Requirements in terms of performance, dynamics, environmental service conditions, physical characteristics, etc. are provided in the appendices of SS-US-200, GPS User System Segment Phase II Specification. To the extent possible, the contractors will also be provided with performance levels specified against mission requirements and with mission scenarios to better understand these requirements.

---

* Private Communication, Mr. B. Glazer, Magnavox Advanced Products Division

Other requirements relate to coordinate references and interoperability. The common global coordinate system for GPS is the World Geodetic System (currently, WGS-72). This is an earth centered, earth fixed coordinate system and is the basis for all GPS calculations. This coordinate reference is convertible to other references such as the Military Grid Reference Systems (MGRS) and to local datums. The UE developed must be capable of coordinate conversion, if required.

GPS user equipment must be capable of being integrated with Inertial Navigation Systems (INS), Inertial Measurement Units (IMU), Attitude/Heading Reference Systems (AHRS), doppler systems, altimeters, compasses and other host vehicle auxiliary sensors as discussed in Section 6. Additionally, the equipment must be capable of interoperability with other communication, navigation, and identification equipment, such as Joint Tactical Information Distribution System (JTIDS) and Naval Tactical Data System (NTDS).

4.0 DESIGN TO COST/LIFE COST

4.1 DTC Concept

Central to the GPS development sequence is proper balance of expense in the system as a whole[1]. Consequently, the DTC concept was adopted as a management tool very early in the development cycle, and it matured to life cycle cost management in recognition that all cost elements (not just acquisition costs but operations and support costs) were essential to making selections and decisions. DTC was judged applicable to GPS as a whole and the user equipment in particular because:

a. The program was then in the conceptual stage, thereby allowing the DTC goals be established early in the development cycle;

b. The program had large potential production and the potential for cost escalation if not controlled;

c. No new technology was involved and it was a straight engineering development program;

d. With DTC goals firmly established, it represented an opportunity for meaningful competitive procurement; and

e. The program is cost effective.

4.2 LCC Management

The design to cost or design to life cycle cost is a cost-control management concept, the principal objective being to introduce a cost target which considers the cost of acquisition and ownership to be achieved by practical tradeoffs between operational capability, performance, cost and schedule. Here the performance characteristics selected for trade-offs against the cost are generally desired but not minimum performance requirements. Cost, as a key design parameter, is addressed on a continuing basis and as an inherent part of the development and production process.

Because of the large population of users and their environmental performance, interface/integration and logistics requirements, the cost of user equipment becomes the largest single item of cost in GPS implementation. For cost effectiveness, past system-level trades conducted during system concept formulation favored moving system complexity from the user equipment toward the space vehicle and in turn toward the control system. That is why greater design-to-cost attention has been focused on the User Segment than the other two.

4.3 Competitive Development

Another fundamental development guideline carried forward since system concept formulation stage is the concept of competitive development. Competition provides a motivation for contractors to exert their best ingenuity in design to keep the cost down. In GPS UE development, competition is encouraged through parallel development of the fundamental configurations in duplicate by independent contractors. Close track and cross check of cost data at any stage through the utilization of DTC/LCC techniques and management provide a safeguard against the danger of contractors giving optimistic quotations which lead ultimately to over runs once the competition has stopped. Of course, at this stage, the option of running the competition far into the production phase is still open.

4.4 DTC/LCC Objectives and Tasks

The Phase II DTC/LCC program objectives are: (1) support the development of a family of GPS user sets at minimum life cycle cost consistent with adequate performance and functional capability; (2) to determine "design to" cost goals, reliability and maintainability goals; and (3) to identify hardware and software support concepts. To achieve these objectives, the contractors are asked to perform trade-offs between operational capability, performance, life cycle cost, and schedule.

The LCC factors such as Unit Production Cost, Initial Production Facility cost, Operation and Support costs and reliability and maintainability requirements consistent with the "design to" goals will be determined by the contractors during Phase IIA through interactive design trades such as those mentioned above to obtain the lowest LCC. The GPS User Equipment Life Cycle Cost Model has been developed. It will be the baseline model to be used in sensitivity analysis and trade studies by the contractors, although some trade studies may require modifications.

(1) Smith, D. L. and Butterfield, F. E., "Navigation Satellite Design for Low User Cost", International Telecommunications Conference, Atlanta GA, Nov. 26-28, 1973.

The key points of the selected DTC approach are: the establishment of LCC as a design parameter; the emphasis on trading off performance characteristics and schedule against cost; and the requirement for continuous evaluation of system development against pre-established cost criteria.

5.0 THE MODULAR DESIGN CONCEPT

5.1 Design Approach

Considering the factors of low LCC requirements, diverse mission applications, and the varied maintenance and logistics strategies among the services, a cost effective approach would be to use a building block approach in a modular fashion in the design and configuration of the user equipment. The idea is that each set will be based on a small number of common hardware and software elements, selected out of a standard GPS collection. The desired result is a small family of electrical, mechanical, and software building blocks from which any user set can be derived, thus promoting the advantages of standardization and high volume production.

To keep the logistics manageable and the equipment nomenclature system workable, the number of unique set designs will be minimized. Each set consists of the hardware and computer programs necessary to convert the GPS navigation signals into timing data, positioning data, navigation data and control and display signals as required. The set associated with each application need not have a unique configuration. The application of a set or sets to more than one application shall be accomplished where technically feasible and economically desirable. The goal is to develop a common module approach to provide a minimum number of differing sets to meet a maximum number of applications.

The hardware elements in a set that are physically separate and distinguishable will be treated as line replaceable units (LRUs). Within the framework of each standard LRU and computer program configuration item (CPCI), the building blocks used to synthesize an LRU or CPCI will come from limited inventory of standard function replaceable hardware and computer programs.

5.2 Required Efforts

The tasks that Phase II contractors face are (1) to consolidate and logically group the requirements which will form the basis for definition of each type user equipment; (2) to perform functional partioning of the set in terms of the functional areas listed in Table III; (3) to implement each functional area with standard replaceable functional modules to synthesize all LRUs, and (4) to define unique set configurations through potential arrangements of the functional areas based on the concepts shown in Figure 1.

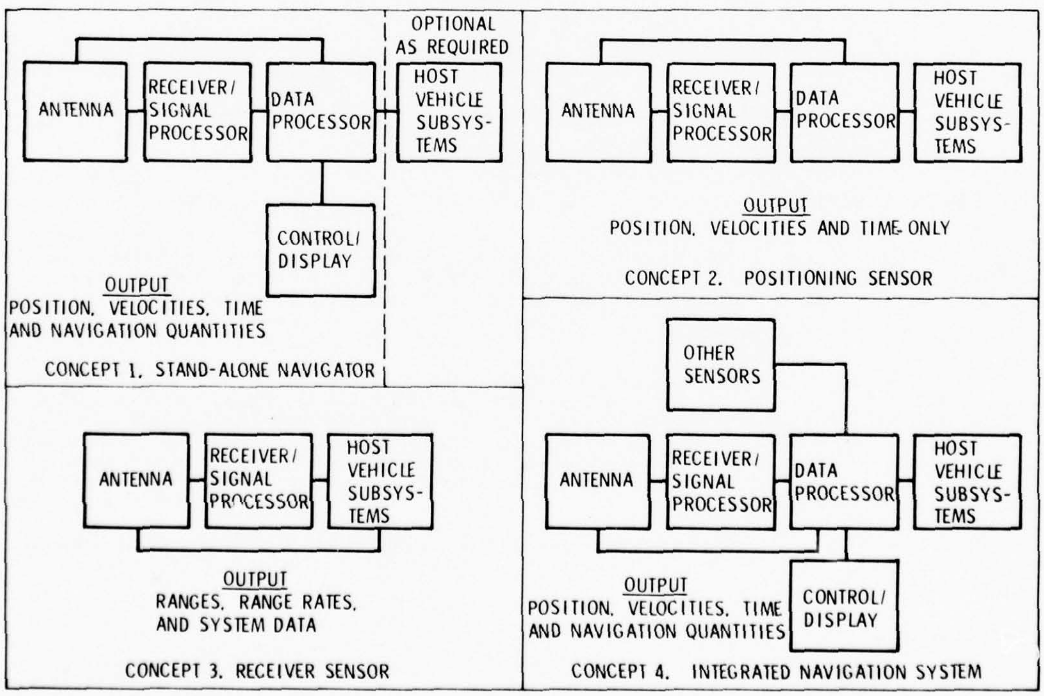


Figure 1. Set Concepts

Table III. Functional Areas

Functional Areas			
1.	Antenna	6.	Power Supply
2.	Receiver/Signal Processor	7.	Flexible Modular Interface
3.	Data Processor	8.	Chassis Components
4.	Computer Programs	9.	Equipment Mounts
5.	Control/Display		

5.3 Functional Concepts

Concept 1. Concept 1 is a stand-alone navigator that uses GPS LRUs exclusively to provide position, velocity, time and derived navigation quantities (e.g., waypoint navigation, steering indications, and approach aids as required) to the operator or host vehicle subsystems in the required format. A minimum amount of aiding may be accepted from, or a minimum amount of information may be exchanged with other host vehicle subsystems depending on the requirements of the particular host vehicle installation.

Concept 2. Concept 2 is a positioning sensor that uses GPS LRUs exclusively to provide position, velocity and time to the operator or host vehicle subsystem in the required format. Information and aiding shall be exchanged with other available host vehicle systems as required.

Concept 3. Concept 3 is a receiver-sensor that provides input quantities (e.g. ranges and range rates) to related host vehicle subsystems where positioning and navigation calculations are performed as required by the missions. Information and aiding shall be exchanged with other available host vehicle subsystems. Prompt aiding (i.e., aiding signals supplied directly to the receiver from external sensors with a minimum of transport delay) will be employed as required.

Concept 4. Concept 4 is an integrated navigation system that uses both GPS-derived and host vehicle subsystem inputs for the navigation process. Concept 4 includes use of host vehicle subsystems to provide aiding information to the GPS receiver to enhance dynamic tracking performance and resistance to jamming.

In keeping with the modular approach to hardware development, the Phase II contractors are asked to develop the UE computer programs, including code classified as firmware, in a modular fashion to facilitate easy adaptation to the various user applications. Design trades of unique software versus common modules must be performed. In order to maximize the utility of these computer program modules, the computer program's dependence on hardware must be minimized.

The importance of software development cannot be overemphasized. Each Phase II contractor is required to prepare a Computer Program Development Plan (CPDP) pertaining to both technical and management of CPCI development. The plan will reflect a carefully thought out approach from initial definition through module partitioning to the delivery of the completed product code.

6.0 INTEGRATION CONCEPT

6.1 Importance of Integration

Besides accomplishing efficient designs of high performance modules useful as prototypes of building blocks to be assembled into Line Replaceable Units to serve the categories of Table II and very likely others (missiles, satellites), it is an objective of the Phase II of the GPS program to make significant progress in the integration of GPS equipment into host vehicle operating systems.

6.2 Phase I

Phase I has accomplished some important steps in this direction: the X-aided equipment, for example, combines a four simultaneous signal receiver with an inertial navigation system according to Figure 2. The computations undertaken involve the Kalman filter states listed in Table IV. Fundamentally, the receiver outputs permit frequent resetting of the INS position. Other computations permit the triggering of a bomb release as a designated target is approached. The equipment is mounted external to a standard F-4 and drives only a pilot's steering indicator in the cockpit and control and display units in the second seat.

The Phase I Z-set is another example of rudimentary integration. This set is required to fit the space occupied by the AN/ARN-118 TACAN receiver/transmitter and is required to interface with apparatus in Figure 3.

6.3 Phase II Integration

In Phase II it is the intention to integrate GPS equipment deeply into some host vehicles for the purpose of obtaining maximum performance on the one hand and for developing integration techniques on the other. One example is shown in Figure 4, the block diagram of the expected F-4G system**. The AN/ARN-101 is a weapon delivery system including an INS which feeds rate information to its radio sensor for the purpose of aiding signal tracking loops. Probably more important is the position, time, velocity and covariance data useful in estimating INS errors. With such data it is possible to obtain high quality INS navigation with simpler and less expensive inertial elements.

** Derived by ARINC Research Company under GPS contract.

Table IV. System State Vector

12 States (X, Y, Z refers to tangent plane coordinates)

$\delta\theta_X$	Angular position error about X axis
$\delta\theta_Y$	Angular position error about Y axis
δH	Altitude error
δV_X	X velocity error
δV_Y	Y velocity error
δV_Z	Z velocity error
$\delta\varphi_O$	User clock phase offset
δF_O	User clock frequency offset
δH_B	Baro altimeter bias error
ψ_X	Platform-to-computer misalignment angle
ψ_Y	Platform-to-computer misalignment angle
ψ_Z	Platform-to-computer misalignment angle

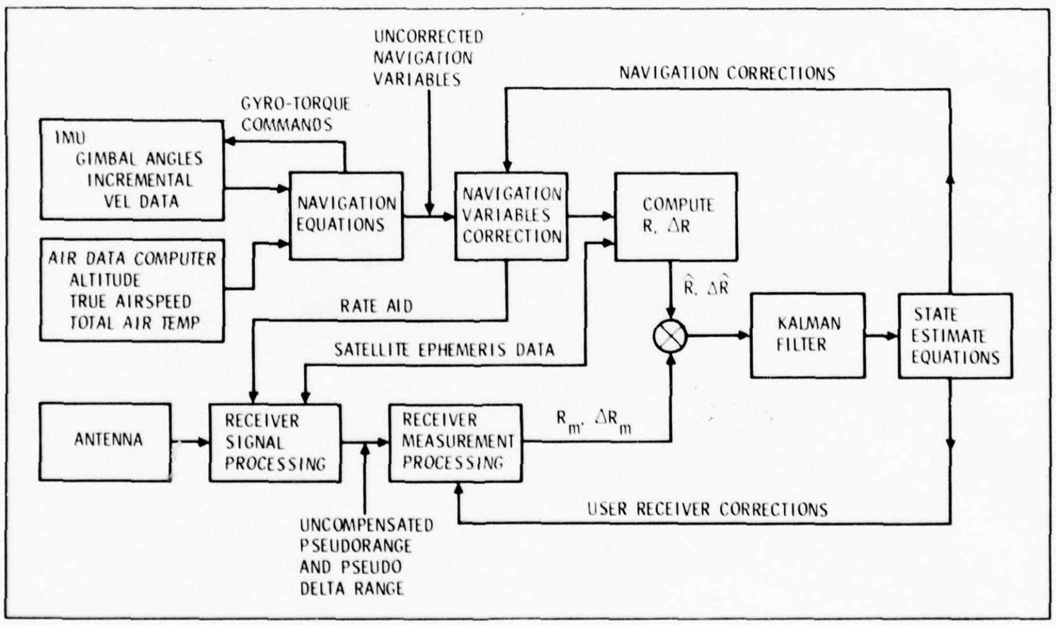


Figure 2. X-Aided Assembly

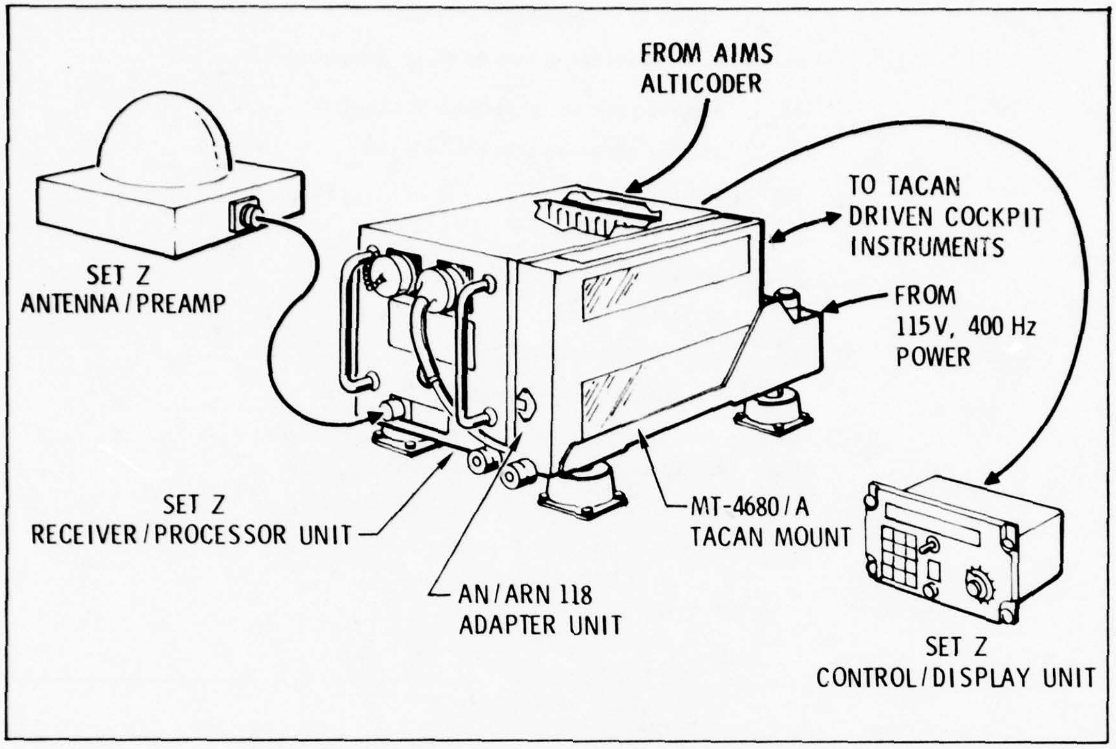


Figure 3. Z-Set Interfaces, Phase I

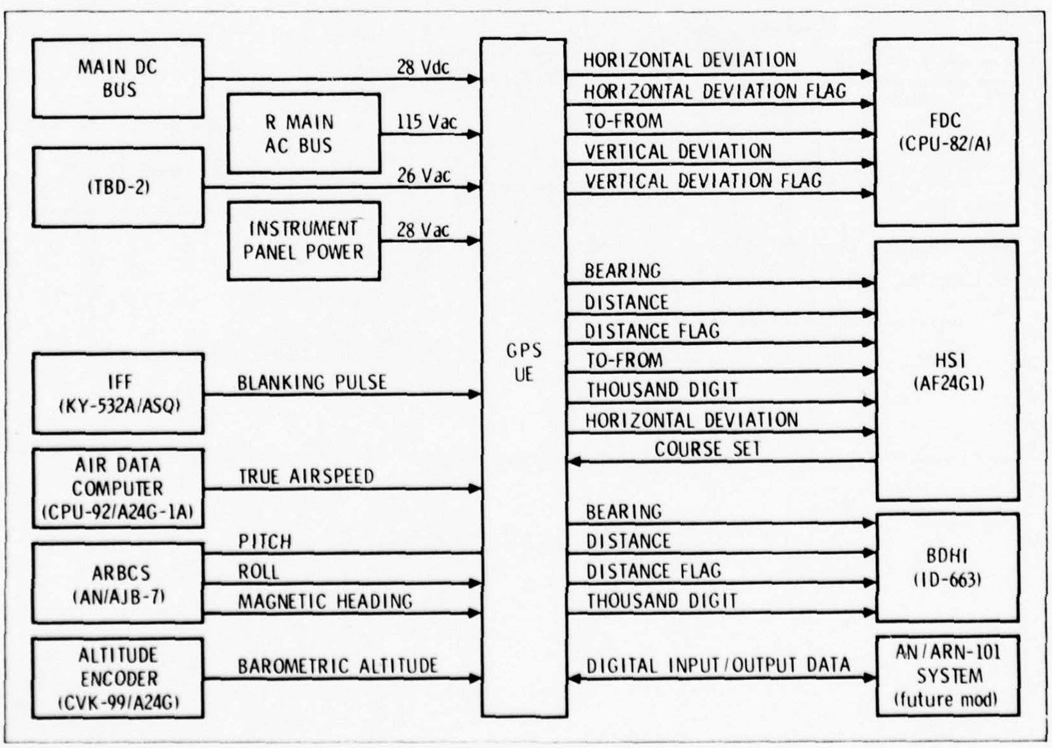
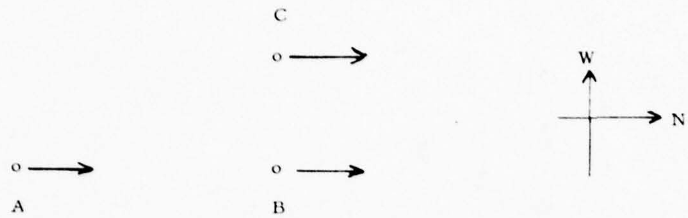


Figure 4. GPS Interfaces to F-4G, Phase II

The ability to restrict INS error by means of radio sensor inputs permits the INS to function as an accurate navigation "flywheel" through areas of intense radio jamming. It is especially important in cases of low rate radio sensor sampling such as is typical of the TDMA systems, where the presence of a reliable dead reckoning device is required to produce even modest relative navigation performance in the presence of navigator dynamics.

The stabilizing or synchronizing of time bases for such operations is also important. Take the simplest case of relative navigation given in the following sketch, where three aircraft set out to fly north disposed as shown.



Since the three sides of this triangle can be measured, its solution is always possible and relative navigation is successful in absence of errors. However, if the time base of B is fast, range C to B (measured at B) will be larger than the true range. B will be inclined to turn west to maintain the intended relationship. The range B to C (measured at C) will be smaller than the true range and C will be inclined to turn west to maintain the intended relationship. A similar effect takes place between A and B causing them to reduce velocity. If A is regarded as "master" and maintains velocity, B interprets his readings as being the consequence of air speed errors and may compensate with a velocity increase. Thus a single synchronization error results in rotation of the configuration and in the linear range results in continual cross track error growth. Redundant information is required to identify this bias and is available from GPS at less cost than that of adding an aircraft to the flight.

Typical interfacing equipments are for the classes of Table II the following:

1. AN/PRC-77; AN/PSC-1; TACFIRE: SINCGARS
2. HR/BDI; HSI; Altimeter; Doppler Radar; Fire Control; Hover Couples
3. INS; Fire Control; Tactical Information Distribution Systems; Flight Indicators
4. Flight Director; INS; Weapon Delivery Systems; Flight Indicators
5. INS; Weapon Delivery Systems; AHRS; Flight Indicators
6. INS; E. M. Logs; Gyrocompass; Doppler Sonar; Fire Control
7. INS; Fire Control; Communications
8. INS; TACAN; ILS Indicators
9. Standard Flight Indicators

6.4 Flexible Modular Interface

While this last list is not all inclusive, it is apparent that the GPS equipment will interconnect to similar things in many of its different employment categories. This leads to the concept of standardizing the GPS outputs as well as the internal circuits. The variety of interfacing requirements is so wide that not all combinations will be accomplished economically in a common interface device, but it is the intent to find one which handles the majority of situations. This "Flexible Modular Interface" is a device which converts navigation data to the proper format for transmission and/or reception between the UE and other host vehicle avionic systems. If special interface devices are necessary, they will be conceived as part of the installation.

6.5 F^3 Specification

All of these considerations, then become elements of the "form, fit, and function" specifications which are the basis for prototype definition. To the degree that they influence the interior of GPS equipment, they must be accommodated. Beyond that, the GPS electronics design details become options for the equipment manufacturers to choose from in minimizing equipment cost.

PERFORMANCE ENHANCEMENT OF THE GPS RECEIVER BY DATA-FREE OPERATION*

J. R. Sklar
L. L. Horowitz

Lincoln Laboratory
Massachusetts Institute of Technology
Lexington, Massachusetts 02173 USA

Introduction

In this section we discuss the operation of a receiver designed to receive the signal transmitted by a Global Positioning System (GPS) satellite in the dynamic environment experienced by a high-performance tactical aircraft. As indicated by the section heading, we concentrate on an operating mode known as "data-free" mode. Because the GPS signal is modulated with system data, the signal energy is spread over a 50 Hz band; as a result, there is a limit to the noise filtering that can be performed. If this data were removed, or (equivalently) stored in the receiver, additional filtering could be employed and other changes made in the receiver to improve the noise (and/or jamming) immunity. The quantitative effects of these changes are derived below.

Data-Free Operation

As already discussed in an earlier section, the signal received from each GPS satellite consists of a Doppler-shifted carrier which is biphase modulated by a known pseudo-random noise sequence and low-rate (50 baud, assumed unknown) data. In the receiving equipment, the carrier and pseudo-random noise code are tracked to provide relative velocity and relative range information, respectively. The low-rate data contains system information such as time-of-day, satellite ephemerides, etc., which the receiving equipment must use to convert the relative velocity and range estimates into absolute velocity and position fixes.

Various forms are possible for the receiver operating in the normal (i.e., not data free) mode. Usually Costas carrier-tracking loops track the Doppler-shifted carrier. This carrier-recovery technique is designed to generate a coherent phase reference that is independent of the binary modulation. Such an approach is necessary in estimating Doppler because of the phase uncertainty that is introduced in the received signal by the low-rate data stream.

The phase uncertainties introduced by the data also affect the choice of code-tracking circuitry. A non-coherent delay-locked loop, also known as an envelope correlation delay-lock discriminator, will track the code and hence estimate relative range without knowledge of the data bits. Since the data is normally unknown at the receiver, the receiver must be designed using such a technique.

Thus, the basic techniques used in both the carrier and code tracking loops are affected by the unknown phase modulation imposed by the data and Doppler velocity uncertainty. However, we may make note of two important considerations:

1. Carrier and code-tracking circuitry that operate with signals of known phase modulation give better performance than equivalent circuitry designed to track signals of unknown phase modulation.

2. In the operational context of the GPS system, the data message is brief and is periodically repeated; further, the data contained in the message changes infrequently, perhaps once every hour or two, and the time-of-day when the message might change can be firmly established.

These two facts provide an opportunity for the GPS receiver to operate as if the data were known. We shall refer to this as "data-free" operation. In this case, the carrier can be tracked with a phase-locked, as opposed to a Costas, loop. As we shall see, the signal-to-noise requirements for a given level of tracking performance, are reduced significantly when this receiver configuration is used. In addition, the phase-locked loop tracks the Doppler frequency directly while the Costas loop tracks at the doubled frequency resulting from a squaring operation. This lowers the lock threshold of the loop significantly.

Ideally, the pseudo-random noise code could also be tracked with a coherent, rather than a non-coherent, delay-locked loop. However, a coherent loop must rely on the phase-locked loop for its phase reference and at the signal-to-noise levels of interest, the phase-locked loop that is estimating Doppler frequency, produces a fairly noisy estimate of phase. Cycle slips are relatively frequent. With such a noisy phase reference, the coherent delay-locked loop performs poorly. Hence the best receiver configuration when the data is known is a non-coherent delay-locked loop (NCDLL) for pseudo-random code tracking and a phase-locked loop (PLL) for carrier-tracking. The PLL also provides a velocity aiding signal to the NCDLL, as will be discussed below.

A further advantage to data-free operation arises because of bandwidth considerations in the code-tracking loop. The first stage in this loop is a correlation mixer which multiples the incoming rf signal by the receiver's estimate of this same signal. When the data is unknown, its modulation spreads the correlator output over a 50 Hz band, limiting the degree to which noise can be predetection filtered. On the other hand, when the data is assumed known, it can be included as part of the pseudo-noise code and eliminated from the predetection circuitry. Thus the predetection filter in the NCDLL can be narrowed from $(50 + f_{s1})$ Hz to $(2f_{s2})$ Hz where f_{s1} and f_{s2} are the frequencies of cycle slips in the Costas loop and the phase-locked loop, respectively.

---

* This work was sponsored by Department of the Air Force.

Operation of the receiver in data-free mode imposes some constraints on the overall GPS system. Principally, the data cannot change unexpectedly. If it does when the receiver is operating in data-free mode, the receiver will undoubtedly lose lock and reacquisition will be necessary. Thus the receiver controller must keep track of the current message, the system time-of-day, and the time of expected message changes, being careful not to operate in data-free mode across a message-to-message boundary.

Analysis

Having described the concept and benefits of data-free operation in general terms, we will next illustrate the reduction in signal-to-noise requirements that can arise from its use. This analysis will be based on the performance of a non-coherent delay-locked loop (NCDLL) that is rate aided by a Doppler-tracking loop. A comparison will be made between the signal-to-noise required for the specified GPS range accuracy under normal operating mode, and the signal-to-noise required for the same accuracy in data-free mode, where a phase-locked loop is used in place of the Costas loop. Reduction in signal-to-noise requirements for other receiver functions are described elsewhere.[1]

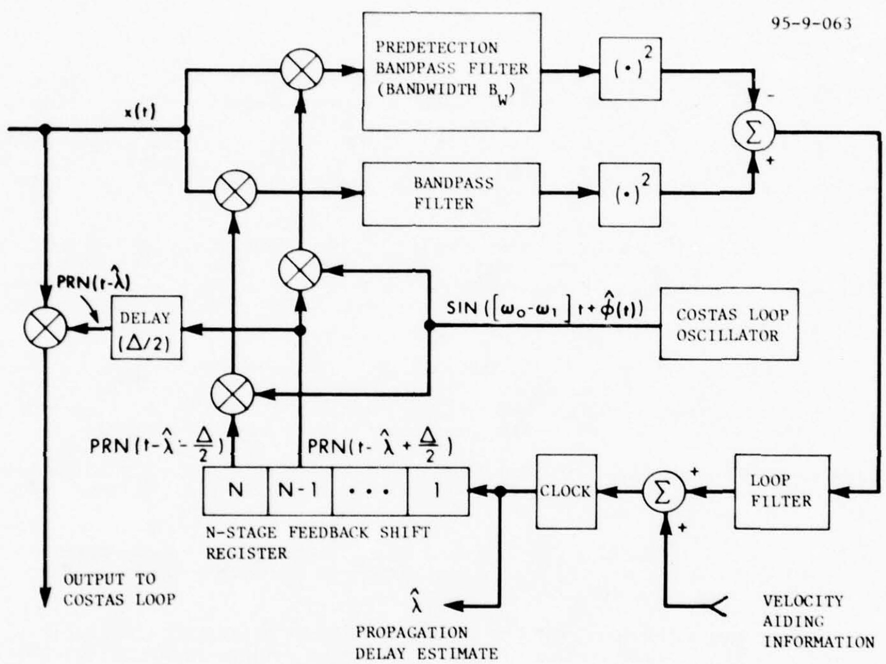


Fig. 1. Diagram of a non-coherent delay-locked loop.

A diagram of the NCDLL is given in Fig. 1. Here the input signal is given by:

$$x(t) = \sqrt{2S} \cdot \text{PRN}(t-\lambda) \cdot \text{data}(t-\lambda) \cdot \sin[\omega_o t + \phi(t)] + n(t) \quad ,$$

where λ is the delay to be estimated by the loop. The feedback shift register generates the pseudo-random noise (PRN) sequence which, together with the data, is modulating the received carrier. We have included a "velocity aiding" input which can include rate information from the carrier tracking loop and/or an inertial measurement unit, if available.

The specified satellite-to-user GPS range measurement accuracy of 1.5 m corresponds to an rms delay error of .05 chips. This delay error, σ_ϵ, is composed of two components: the first, σ_{ϵ_n}, is due to input noise, while the second, σ_{ϵ_d}, is due to user dynamics. Thus, the total rms delay error is given by:

$$\sigma_\epsilon = \left(\sigma_{\epsilon_n}^2 + \sigma_{\epsilon_d}^2\right)^{1/2} \quad . \tag{1}$$

Gill[2] has analyzed the NCDLL and shown that the rms delay jitter caused by input noise, is given by:

$$\frac{\sigma_{\epsilon_n}}{L_c} = \{[(N_{oe}/2S) + (N_{oe}/S)^2 B_W] B_{LC}\}^{1/2} \text{ chips} \quad , \tag{2}$$

where L_c is the chip length (30 m), B_W is the predetection filter bandwidth (50 Hz + f_s) and B_{LC} is the one-sided noise bandwidth of the loop. From equation (2) it is clear that the narrower the loop noise bandwidth, B_{LC}, the better the tracking performance for a fixed S/N_{oe} ratio. However, the loop bandwidth, B_{LC}, is bounded from below by user dynamic motion.

We first analyze the receiver operation in normal (i.e., not data-free) mode. If the NCDLL were operating on its own, with no aiding aside from the carrier frequency estimate from the Costas loop oscillator, the next step in the analysis would involve establishing a model of the user dynamic motion, and determining the loop bandwidth which provides the required σ_ϵ with a minimum (S/N_{oe}). However, we have already stated that a velocity reference signal from the Costas loop can be introduced into the NCDLL at the aiding node. Then, the dynamic motion which the NCDLL must track is effectively the difference between the user dynamics and the reference signal from the Costas loop. In fact, because the optimal Costas loop bandwidth is significantly higher than the NCDLL bandwidth in this configuration, the user dynamics are effectively eliminated by the Costas loop.

On the surface, this would suggest that the loop bandwidth is no longer bounded from below, and could be made as small as desired. However, the Costas loop does not track the carrier perfectly. Cycle slips in the Costas loop aiding signal become the dominant difference between the user dynamics and the reference signal; the NCDLL bandwidth must be wide enough to accommodate them. Thus, we must jointly optimize the Costas loop bandwidth, which determines the cycle slips in the velocity reference, and the NCDLL bandwidth to achieve the specified total delay jitter with the minimum required signal-to-noise ratio.

The Costas loop bandwidth, B_{LP}, the signal-to-noise ratio, S/N_{oe}, and the user dynamics determine the frequency of cycle slips occurring in this loop. The NCDLL must track the cycle slips well enough to prevent excessive delay error and hence ranging error. Before we can analyze the receiver in detail, we must specify the user dynamic environment and develop a model for the dynamic error in a tracking loop caused by these dynamics.

In the high performance tactical aircraft application, one can expect that the user velocity, acceleration, and jerk will have bounds given approximately by:

$$V_{max} \leqslant 900 \text{ m/s}$$

$$A_{max} \leqslant 50 \text{ m/s}^2$$

$$J_{max} \leqslant 100 \text{ m/s}^3 \quad . \tag{3}$$

We will use a sinusoidal model for the user dynamics, where the amplitude of the sinusoid is bounded by Eq. (3). Thus, if the user position sinusoid is to have angular frequency ω, then the user position, velocity, acceleration, and jerk are given by:

$$p(t) \triangleq P \cos(\omega t)$$

$$v(t) = -\omega P \sin(\omega t)$$

$$a(t) = -\omega^2 P \cos(\omega t)$$

$$j(t) = \omega^3 P \sin(\omega t) \quad . \tag{4}$$

The bounds in Eq. (3) may be applied to Eq. (4), with the resulting bound on P [defined in Eq. (4)] as a function of ω.

$$P_{max}(\omega) = \min\,[900/\omega,\ 50/\omega^2,\ 100/\omega^3]\ \text{(in meters)} \quad . \tag{5}$$

Costas phase-tracking loops with linearized third-order transfer functions are chosen in order to meet the criterion that there be no steady-state error to user accelerations. The closed-loop response of an optimized third-order loop is given by.[3]

$$H(s, B_L) = \left[\frac{(\frac{12}{5}) B_L s^2 + (\frac{72}{25}) B_L^2 s + (\frac{216}{125}) B_L^3}{s^3 + (\frac{12}{5}) B_L s^2 + (\frac{72}{25}) B_L^2 s + (\frac{216}{125}) B_L^3}\right] \quad , \tag{6}$$

where B_L is the one-sided noise bandwidth of the loop.

This response relates input phase (i.e., position) to output phase (i.e., position). The transfer function from the loop input to the error node is given by:

$$[1 - H(s, B_L)] = \left[\frac{s^3}{s^3 + (\frac{12}{5}) B_L s^2 + (\frac{72}{25}) B_L^2 s + (\frac{216}{125}) B_L^3}\right] \quad . \tag{7}$$

The peak signal at the error node, for a sinusoidal user position input, $P \cdot \cos(\omega t)$, is given by:

$$e_{d_{max}}(\omega, B_L) = P \cdot |1 - H(j\omega, B_L)| \quad . \tag{8}$$

Thus, using Eq. (5) we may obtain the peak signal at the error node with the specified dynamics, as a function of frequency.

$$e_{d_{max}}(\omega, B_L) = P_{max}(\omega) \cdot |1 - H(j\omega, B_L)| \quad . \tag{9}$$

Hence, the peak error with the given loop bandwidth may be found as the maximum of Eq. (9) over all ω.

$$e_{d_{max}}(B_L) = \max_{\omega} \left[e_{d_{max}}(\omega, B_L) \right] = \max_{\omega} [P_{max}(\omega) \cdot |1 - H(j\omega, B_L)|] \quad . \tag{10}$$

Equations (5) and (7) are substituted into Eq. (9) and computer plots are made of $[e_{d_{max}}(\omega, B_L)]$ for various ω and B_L. This dynamic error may be converted to radians (33.5 radians/m at 1600 MHz) to obtain $\sigma_{\phi_d}^{(c)}$, the phase error due to user dynamics in the Costas loop. We find, then, that

$$\left\{ \begin{array}{llr} \sigma_{\phi_d}^{(c)} = 1940\, B_{LP}^{-3} & \text{radians} \quad ; & \infty > B_{LP} > 2.3 \text{ Hz} \\ \sigma_{\phi_d}^{(c)} = 838\, B_{LP}^{-2} & \text{radians} \quad ; & 2.3 \text{ Hz} \geqslant B_{LP} > .056 \text{ Hz} \end{array} \right\} \quad . \tag{11}$$

Note that the peak value of the sinusoidal error is used instead of the rms value, because the user sinusoidal dynamics may have quite low frequencies.

Viterbi[4] has shown that for a first-order coherent phase-locked loop, with no input dynamics, the frequency of cycle slips is given by:

$$f_s = \frac{AK}{2\pi^2 \alpha I_o^2(\alpha)} \quad , \tag{12}$$

where AK is the gain of the proportional control term in the loop filter of the phase-locked loop, I_o is the zero-th order modified Bessel function, and

$$\alpha \triangleq 1/[\sigma_{\phi}^{(p)}]^2 \quad , \tag{13}$$

where $\sigma_{\phi}^{(p)}$ is the rms phase jitter calculated from a linearized phase-locked loop model. By analogy, for the Costas loop (including input dynamics and the factor for the doubled frequency) we apply the same result using

$$\alpha' \triangleq 1/[\sigma_{\phi}^{(c)}]^2 \quad , \tag{14}$$

where $\sigma_{\phi}^{(c)}$ is the rms phase jitter of the Costas loop. This jitter is due in part to the dynamic loading, Eq. (11), and in part to the noise,[4] as given below.

$$\sigma_{\phi_n}^{(c)} = \{[(N_{oe}/S_1) + (N_{oe}/S_1)^2 (B_w/2)]\, B_{LP}\}^{1/2} \quad \text{radians} \tag{15}$$

where S_1 is the effective input signal power to the Costas loop, which is a function of the delay jitter of the delay-locked loop $[S_1 \simeq (1 - \sigma_\epsilon)^2 S]$, B_w is the predetection filter bandwidth, and B_{LP} is the one-sided noise bandwidth of the Costas loop.

Combining the dynamic and noise contributions and applying Eq. (14), we obtain

$$\alpha' = .25\, \{[(N_{oe}/S_1) + (B_w/2)\, (N_{oe}/S_1)^2]\, B_{LP} + 3.764 \times 10^6\, B_{LP}^{-6}\}^{-1} \quad . \tag{16}$$

Now, since from Eq. (6),

$$AK = \left(\frac{12\, B_{LP}}{5}\right) \quad , \tag{17}$$

we obtain the desired expression for f_s

$$f_s = \left[\frac{12\, B_{LP}}{10\pi^2 \alpha' I_o^2(\alpha')}\right] \quad , \tag{18}$$

which is a function of S/N_{oe} and B_{LP}. In fact, for a given S/N_{oe} we may minimize f_s by proper choice of B_{LP}.

The NCDLL must track the cycle slips from the Costas loop in order that the dynamic loop error not be beyond specification. The cycle slips represent steps in delay to the NCDLL. Assume, as a worst case, that the cycle slips are all in the same direction for some period of time. The rise time of the NCDLL is given approximately by $(1/2\, B_{LC})$, where B_{LC} is the one-sided code loop bandwidth. Thus the loop error will be the total delay generated by half the cycle slips which occur in time $(1/2\, B_{LC})$. Hence, the error is that represented by $(f_s/4\, B_{LC})$ slips. Each slip represents half a cycle at 1600 MHz (the

Costas loop tracks a double frequency). This half cycle is .094 m, or .0031 P-code chips. Thus, the code loop tracking error is given by:

$$\epsilon_d = \left[\frac{(.0031)\ f_s}{4\ B_{LC}}\right] \quad \text{chips} \quad . \tag{19}$$

Substituting Eq. (19) and Eq. (2) into Eq. (1), we obtain:

$$\sigma_\epsilon = \left\{\left[\frac{(.0031)\ f_s}{4\ B_{LC}}\right]^2 + [(N_{oe}/2S) + (50 + f_s)\ (N_{oe}/S)^2]\ B_{LC}\right\}^{1/2} \quad \text{chips} \quad . \tag{20}$$

We set $B_W = 50 + f_s$ to allow for cycle slips in the predetection filter bandwidth.

The required S/N_{oe} for a given degree of code tracking can be determined as follows. A trial value of (S_1/N_{oe}) is substituted into Eq. (18), and B_{LP} is chosen to minimize f_s. B_W is set to 50 Hz for the first iteration of this calculation. The obtained value of f_s is substituted into Eq. (20), and B_{LC} is chosen to minimize the (S/N_{oe}) required for the desired value of σ_ϵ. The initial trial value of (S_1/N_{oe}) is adjusted using the obtained value of (S/N_{oe}). The iteration is continued until (S/N_{oe}) converges.

If this procedure is followed for the basic GPS receiver, not operating in data-free mode, we can calculate the S/N_{oe} requirement for maintaining code accuracy by setting Eq. (20) equal to .05 chips, which corresponds to the specified ranging accuracy of 1.5 m (10 MHz = 30 m). The result of the iterative solution is S/N_{oe} = 19 dB · Hz, with B_{LP} = 18 Hz and B_{LC} = 0.093 Hz. Similarly we can obtain the operating point for code lock using σ_ϵ = .3 chips, obtaining S/N_{oe} = 13 dB · Hz, B_{LP} = 11 Hz and B_{LC} = 0.20 Hz.

We have already qualitatively described the basis for improved performance in data-free mode. To make a quantitative assessment we must determine the effects of the receiver changes that become possible in this mode.

First, we have explained earlier that a phase-locked loop can be used for carrier tracking instead of a Costas loop. Since the rms phase jitter at the output of a phase-locked loop, caused by input noise, is[4]

$$\sigma_{\phi_n}^{(p)} = \left(\frac{N_{oe}\ B_{LP}}{S_1}\right)^{1/2} \quad \text{radians} \quad , \tag{21}$$

we use this expression in place of Eq. (15), in an equivalent analysis to the one above. Note that the second term in the brackets of Eq. (15) is no longer present. Since this term is comparatively large when S_1/N_{oe} is small, this change to the receiver has a significant effect on the signal-to-noise requirements. In addition, the phase-locked loop tracks a single, as opposed to a double, frequency; as a result the lock threshold is significantly lowered (the time jitter can be twice as large before a cycle slip occurs).

Second, the predetection bandwidth of the code loop can be narrowed from $(B_W = 50 + f_{s_1})$ to $(B_W = 2 f_{s_2})$ where f_{s_1} and f_{s_2} are the frequencies of cycle slips in the Costas loop and phase-locked loop, respectively.

If the analysis above is repeated with these modifications for the data-free mode, we conclude that for code accuracy we require S/N_{oe} = 14 dB · Hz, B_{LP} = 16 Hz, and B_{LC} = 0.057 Hz. For code lock, S/N_{oe} = 7 dB · Hz, B_{LP} = 41 Hz, and B_{LC} = .079 Hz. Thus there is a 5 dB reduction in the signal-to-noise required to meet the ranging accuracy specification, and a 6 dB reduction for code lock.

It is also important to comment on the equivalent improvements that can be obtained when data-free operation is coupled with velocity-aiding by an inertial-measurement unit (IMU). If an IMU with an accelerometer scale factor error of .002 is used, the dynamics which the carrier loop must track are effectively reduced by a factor of .002, allowing narrower loop bandwidth and higher noise tolerance. The carrier tracking is then improved and the carrier loop can then aid the code loop at lower signal-to-noise levels.

Quantitatively, the noise immunity obtained with IMU aiding, both in and out of data-free mode, is shown in Table I (together with the previously described results). In fact, this table can serve as a summary of all results developed here. It is clear that data-free operation significantly reduces the signal-to-noise requirements whenever it is used.

TABLE I

SUMMARY OF RECEIVER REQUIREMENTS

Receiver Operating Conditions		Receiver Requirements				
Data-Free Operation (Yes or No)	IMU Aiding (Yes or No)	Item	Range		Range Rate	
			Lock	Track	Lock	Track
No	No	S/N_{oe} (dB · Hz)	13	19	27	33
		B_{LP} (Hz)	11	18	27	33
		B_{LC} (Hz)	.20	.093		
Yes	No	S/N_{oe} (dB · Hz)	7	14	20	33
		B_{LP} (Hz)	11	16	22	33
		B_{LC} (Hz)	.079	.057		
No	Yes	S/N_{oe} (dB · Hz)	7	13	19	24
		B_{LP} (Hz)	1.0	2.1	3.5	4.2
		B_{LC} (Hz)	.017	.010		
Yes	Yes	S/N_{oe} (dB · Hz)	−3	4	11	24
		B_{LP} (Hz)	1.0	1.7	2.7	4.2
		B_{LC} (Hz)	.0074	.0073		

REFERENCES

1. L. L. Horowitz and J. R. Sklar, "ECM Vulnerability of the GPS User Receiver in a Tactical Environment," Lincoln Laboratory Project Report XR-1 (7 May 1976).

2. W. J. Gill, "A Comparison of Binary Delay-Lock Tracking-Loop Implementations," I.E.E.E. Trans. on Aerospace and Electronic Systems, Vol. AES-2, No. 4 (July 1966) pp. 415-424.

3. R. Jaffe and E. Rechtin, "Design and Performance of Phase-Lock Circuits Capable of Near-Optimum Performance Over a Wide Range of Input Signal and Noise Levels," I.E.E.E. Trans. on Information Theory, Vol. IT-1 (March 1955) pp. 66-76.

4. A. J. Viterbi, Principles of Coherent Communication, McGraw-Hill, New York (1966).

INTEGRATION OF GPS WITH INERTIAL NAVIGATION SYSTEMS

by

Duncan B. Cox, Jr.
The Charles Stark Draper Laboratory, Inc.
555 Technology Square
Cambridge, Massachusetts 02139

SUMMARY

GPS navigation systems can be utilized for extremely accurate navigation without the use of auxiliary sensors. However, when the vehicle being navigated is undergoing dynamic rotation and/or translation, an inertial navigation system can be combined with a GPS navigation system to provide improvements in navigation performance and antijam performance over what could be achieved with either system alone. This paper describes various techniques that can be utilized in the design of an integrated GPS-inertial navigator to achieve these improvements, particularly for applications to tactical fighter aircraft. The degress of implementation complexity of the techniques and the degrees of improvement in performance are indicated qualitatively. Some of the techniques are new and not yet fully developed, but may prove useful in future applications where performance is particularly important.

1. INTRODUCTION

The state of the art in inertial navigation has been under development for over two decades, is widely utilized in military and civilian applications, and is well documented in the technical literature. On the other hand, the NAVSTAR Global Positioning System (GPS) has been under development for a much shorter period and has not yet been utilized to a significant extent in practical applications. With the publication of this AGARDOGRAPH, and a forthcoming special issue of NAVIGATION[1], and with reference to a host of government reports, conference papers, and journal articles, the state of current GPS technology will be quite well documented. Because of the many desirable features of GPS, it is very likely to be utilized world-wide in a very large number of practical applications[2].

In some applications, e.g., for tactical aircraft, the combination of an inertial navigation system (INS) and the GPS offers particular advantages, and integrated GPS-INS systems are being developed to capitalize on these advantages. Early integrated systems will be achieved by integrating GPS navigators with existing INSs. These integrated systems will exploit the synergism between the GPS and inertial subsystems, but because these INSs have not been designed for integration with GPS hardware, exploitation may be only partial. Eventually, GPS subsystems and INS subsystems will be designed for optimal integration and the synergism will be realized more completely in the integrated systems.

Integration of GPS and INS subsystems with other subsystems such as JTIDS[3], will also occur in the future. Although more and more complex functions will be able to be optimized via future digital computing power, it is likely that the major hardware and software subsystems on tactical aircraft will remain substantially autonomous. Maintaining subsystem autonomy will take advantage of the enhanced reliability and maintainability associated with such architectures and the steady reductions in costs of local processors and multiplexors[4].

This paper addresses the techniques and benefits of functional integration of GPS and inertial subsystems. The discussion is primarily in the context of applications to tactical fighter aircraft, where INSs are widely employed already, but the principles discussed are valid for a wide range of applications. References are provided for background information.

2. BENEFITS OF INTEGRATION

GPS is expected to operate in a stand-alone configuration in a benign environment. However, GPS and inertial navigation systems have complimentary features which can be exploited in an integrated system to create synergistic improvements in navigation performance. The improvements are most pronounced when the GPS signal-to-noise ratios are low and the vehicle is undergoing high-dynamic maneuvers.

As long as vehicle dynamics or noise-to-signal ratios are not excessive, the GPS receiver can provide pseudorange and pseudorange-difference data which can be used for estimating errors in position and velocity, and certain other error parameters of the INS and of the GPS receiver clock. The estimates of INS error parameters allow INS navigation with substantially smaller errors than could be achieved with either a GPS navigator or an INS alone. On the other hand, the INS is able to provide accurate "aiding" data on short-term vehicle dynamics to the GPS receiver. By utilizing those aiding signals to effectively reduce the dynamics of the signals to be tracked, the GPS receiver can maintain relatively low tracking bandwidths and can withstand relatively high noise-to-signal ratios, perhaps due to high levels of jamming. The INS signals can

also be used as a basis for pointing narrow-beam antenna patterns at the GPS satellites, thereby also reducing the effects of jamming. When noise-to-signal ratios become so high that tracking of GPS signals is impossible, the INS is capable of navigating independently. The accuracy of independent INS navigation, and also the accuracies of the aiding signals mentioned above, are enhanced by the use of the previously derived estimates of INS error parameters. When GPS signal conditions improve sufficiently to allow tracking, the INS provides data on initial position, velocity, and acceleration for use in reacquiring the GPS codes and carriers quickly. The INS also provides data for use in adapting the tracking-loop parameters to varying conditions of signal dynamics and signal-to-noise ratios, thereby improving the ability of the tracking loops to acquire and maintain lock on the GPS signals. In this fashion the GPS and INS systems can be made to help each other to achieve better performance than could be achieved with either system alone.

3. DESCRIPTION OF STAND-ALONE GPS AND INERTIAL SYSTEMS

Because a variety of stand-alone INSs are currently in use and some GPS receivers are being configured for stand-alone operation, it is instructive to examine briefly these stand-alone systems in preparation for showing how systems with varying degrees of integration might logically evolve from them.

Figure 1 shows several important functional elements of stand-alone GPS and inertial navigation systems, and of some ancillary systems typically appearing on tactical fighter aircraft.

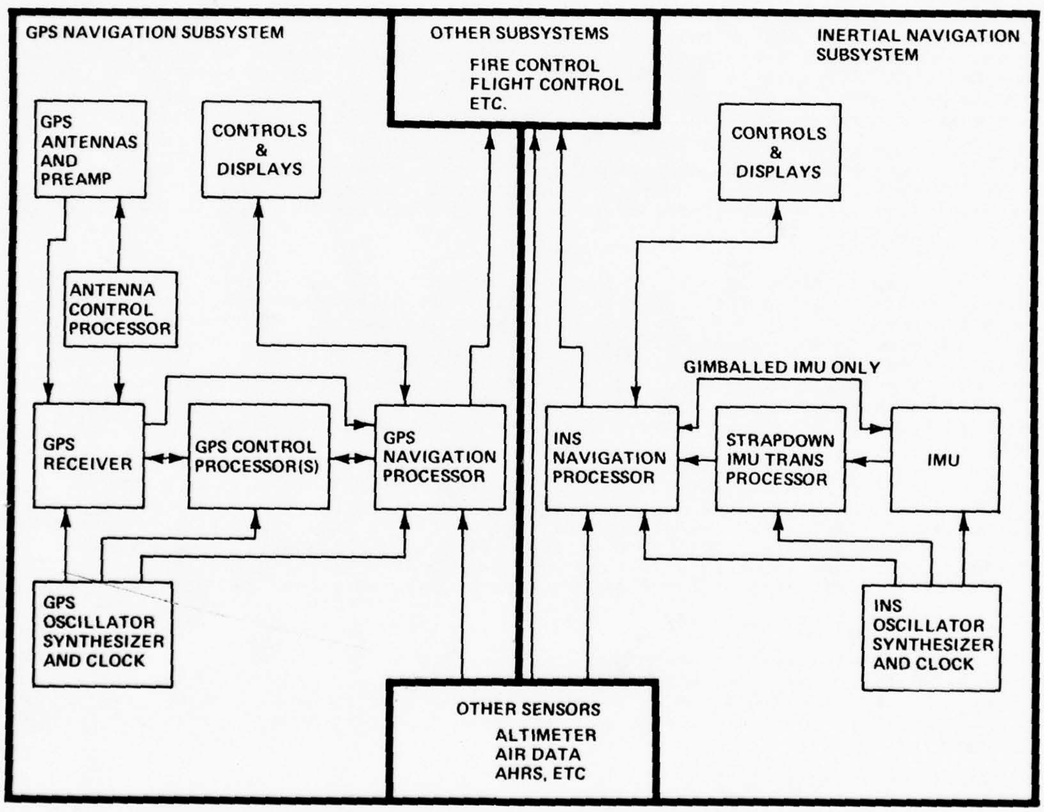


Figure 1. Stand-alone GPS and inertial navigation systems.

The inertial measurement unit (IMU) provides from its accelerometers three axes of time-integrated acceleration (specific-force) information in a reference (usually local-level, wander-azimuth) coordinate frame which is stabilized by its gyroscopes (gyros). Current tactical aircraft utilize gimballed platforms wherein the accelerometers and gyros are mounted on a stable platform, oriented accurately by gimbal-control loops on the basis of information supplied by the gyros.

An alternative, "strapdown", mechanization is also possible, wherein the accelerometers and gyros are mounted on a nonstabilized structure attached to the frame of the aircraft. Data from the strapdown gyros are used for determining the orientations of the accelerometers as they vary with aircraft attitude. The IMU transformation processor

shown in Figure 1 resolves at high speed the strapdown accelerometer data into the gyro-stabilized coordinate system, and is not present in a stabilized-platform mechanization. As inertial components are developed that can provide accurate long-term navigation in the face of the dynamics of the aircraft's attitude, strapdown IMUs are likely to see increasing utilization, and they are candidates for future integration with GPS equipment.

The stabilized accelerometer data are supplied to the navigation processor. Data from an altimeter and from other auxilliary sensors are also supplied. INS mechanization algorithms in the navigation processor propagate the standard inertial navigation equations at a high rate (10 - 50 Hz) and supply torquing commands to the gyros (in the stable-platform mechanization). The INS mechanization algorithms also accomplish gyrocompassing for preflight alignment of the inertially stabilized frame[5-15, 17].

The INS navigation processor may also incorporate Kalman (or less optimal) filter algorithms for improved in-flight estimation of such variables as position, velocity, altimeter bias, misalignments of stabilized axes, and gyro drift-rate biases, on the basis of data from external sources. Examples of such external data are visual position fixes by the pilot, radar fixes, Loran fixes, Doppler radar data, etc. The Kalman filter algorithms require substantial execution time, but because the errors of the unaided INS grow slowly in comparison to the vehicle dynamics the Kalman filter corrections need be incorporated only infrequently. The design of the navigation filter algorithms and the performance of INSs that are aided by auxiliary sensor data have been reported extensively in the literature[11-32]. The INS mechanization and filter algorithms may be incorporated in a single navigation processor as suggested by Figure 1, or in separate processors.

The main functional elements of a possible stand-alone GPS navigation subsystem are shown on the left-hand side of Figure 1. The heart of the subsystem is the GPS receiver which amplifies and filters the GPS radio signals, correlation-detects and tracks the carriers and codes, and correlation-detects the GPS satellite data modulations. The GPS control processor controls the moding of the receiver so as to accomplish acquisition, tracking, etc., and generates estimates of certain signal-processing parameters. The antennas feeding the receiver can be simple ones with broad beam patterns covering the different regions of the space around the aircraft, or they can be capable of beam steering or null steering, in which case an antenna control processor is needed.

The GPS control processor assembles the pseudorange and delta-pseudorange data from the receiver and passes it on to the GPS navigation processor. The navigation processor receives the GPS data and auxiliary sensor data and provides extended Kalman filter estimates of, e.g., position, velocity, and acceleration of the aircraft, time offset and frequency offset of the GPS user clock, and altimeter bias. The unaided (stand-alone) GPS navigation filter must incorporate new GPS data somewhat more frequently than the aforementioned INS Kalman filter incorporates INS data because the aircraft dynamics, rather than the relatively mild INS-error dynamics, must be modeled in the unaided GPS navigation filter. All measurements must be accurately time tagged, particularly if measurements to different satellites are made sequentially, rather than simultaneously. The navigation solution is delivered to pilot displays as well as to other subsystems such as fire control and flight control.

A new GPS navigation solution may be desired as often as every 100 ms, but the complete solution of one cycle of an 11-state extended Kalman filter by the GPS navigation processor generally requires substantially more time, on the order of 5 seconds for one existing design. In that design the covariance matrix and the gain matrix are computed every several seconds and the last gain matrix is used for incorporation of new data more rapidly, about every 300 ms, until a new gain matrix is computed[33,34]. The result is a reasonable compromise between processor cost and navigation performance. In addition, the design provides a unified procedure for incorporating new GPS data from the different tracking loops and for accounting for occasional lapses in data. Such lapses are expected to occur whenever aircraft orientations are unfavorable for reception of satellite signals.

A prime function of the unaided GPS Kalman filter is to incorporate more than the minimum GPS data required for a navigation fix and thereby to minimize the effects of random measurement errors. But many of the measurement errors are correlated in time, and the effects of severe aircraft dynamics tend to dwarf the effects of the random error components. It is possible that simpler deterministic solutions for position and time and for velocity and frequency from the pseudorange and pseudorange-difference data may prove to be advantageous in some applications where speed of response or computational simplicity are prime considerations[35,36]. But algorithms for propagating the navigation solution through occasional lapses in GPS data must still be provided. Regardless of the computational techniques employed, the unaided GPS navigation solution under high-dynamic conditions will always exhibit very large velocity errors in comparison to what can be achieved with an integrated GPS-inertial system.

The operation of the GPS receiver is described in detail in accompanying articles in this AGARDOGRAPH and in other references(37-40).

In the remainder of this paper, various means of integrating the GPS and inertial subsystems in order to obtain improved performance will be discussed. Integration of the navigation filters will be discussed first. Then the discussion will cover aided acquisition, aided tracking, and adaptive tracking. Finally, antenna control will be described briefly. Each of these functional integrations requires implementation of data transfer and/or data processing functions. These additional functions and the structure of an integrated system are illustrated in Figure 2. This figure gives an indication of the degree of increase in complexity associated with the integration functions that are designed to provide increased performance. The increase in complexity is offset to a degree by the opportunity to consolidate navigation filters, and common functions such as controls and displays. It should be stressed that the system should be capable of reconfiguration to stand-alone GPS or inertial navigation systems in the event of failure of key GPS or INS components.

4. INTEGRATED GPS-INERTIAL NAVIGATION FILTERS

The simplest way to integrate the stand-alone GPS and inertial subsystems is to provide a channel for delivering the position and velocity data from the GPS navigation solution to the INS navigation processor. The INS software could simply reset its mechanization algorithms to the positions and velocities specified by the GPS navigation processor. However, this procedure suffers from the relatively large errors made by the unaided GPS navigation filter in the presence of severe aircraft dynamics, from the failure to estimate several observable IMU error parameters, and from the lack of resolution of the typical INS mechanization algorithms.

Substantially better navigation accuracy can be obtained by supplying, with proper GPS-INS time tags, raw GPS data and data from the INS mechanization algorithms (together with auxiliary sensor data) to an integrated GPS-inertial navigation filter as illustrated in Figure 2. The integrated navigation filter supplies Kalman estimates of a combined set of GPS-INS error parameters, e.g., 3 positions, 3 velocities, 3 misalignments of the inertially stabilized frame, 3 gyro drift rates, 1 time, 1 frequency, 1 altimeter bias, and 1 altimeter scale factor. These estimates of error parameters are used as infrequent updates to the INS mechanization algorithms, which in turn provide frequent estimates of position, velocity, and acceleration of the vehicle on the basis of data from the IMU. The number of error states to be mechanized is dependent upon the desired error-update rate, the processing power available, and the number of significant error sources to be encountered. Because the INS system errors grow so slowly, in comparison with the rate of growth of uncertainties in vehicle dynamics without the benefit of IMU measurements, the update rate of the GPS-INS navigation Kalman filter can be much slower than that for the unaided GPS navigation filter while still obtaining substantially superior results (particularly in estimation of velocity). However, there is still the problem associated with execution of the high-order Kalman algorithms at the desired rate. A variety of approximation techniques are employed to ease the computation burden.

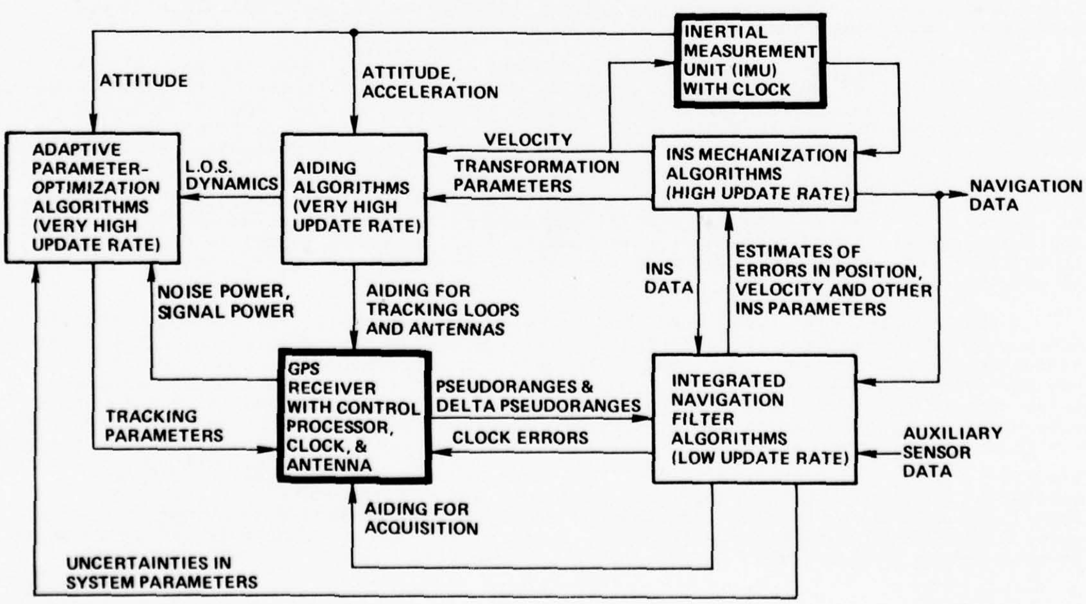


Figure 2. Structure for integrated GPS inertial navigation system.

20-8

The INS mechanization equations, updated by the GPS-INS navigation filter, can be used for accurate INS navigation during periods when GPS is unavailable. The in-flight estimation of IMU alignment and other INS error parameters made possible by the earlier availability of GPS improves the accuracy of INS navigation when GPS is not available. The quality of the IMU must be sufficient to meet the requirements of stand-alone INS navigation under these circumstances.

For best performance, the integrated GPS-inertial system should utilize INS data with resolutions consistent with GPS accuracies, i.e., finer than the resolutions of data from typical stand-alone INSs. The needed increase in resolution can be achieved through modification of the INS mechanization algorithms or by utilizing raw IMU data in separate high-resolution algorithms for inertial extrapolation between GPS updates.

The integrated system navigation filter can be implemented in the GPS navigation processor or in the INS navigation processor of Figure 1. Consideration should be given to allowing the configuration to revert to a stand-alone navigator with only GPS or only INS inputs (together with other auxiliary sensor inputs), in case there is a failure in the GPS or INS equipment. Because of continued progress in increasing the reliability of digital processors, it makes sense in some applications to implement the GPS, INS, and integrated GPS-INS Kalman navigation algorithms in a single processor as suggested by Figure 2. This option should be considered when designing a system for a new aircraft where the INS is not already resident.

5. AIDED GPS ACQUISITION

When the GPS subsystem is jammed, or otherwise inactive, navigation can proceed on the basis of the INS (and auxiliary aids) only. The INS will perform well because its error parameters will have been previously calibrated in flight when the GPS subsystem was operating. Nonetheless, the unaided INS errors will gradually grow, and GPS updates should be used as soon as they are available. When the GPS signals again become usable the INS navigation solution can act as an aid to direct reacquisition of the GPS codes and carriers. The time-tagged INS navigation solution and covariance estimates together with estimates of clock errors and covariances, are used to generate estimates of pseudorange, pseudorange rate, and search regions in time and frequency. In this way the INS not only provides navigation during outages but also greatly shortens the time required for reacquiring the GPS signals after they again become usable. The improved calibration of the GPS clock made possible by inertially aiding the GPS navigation solution also supports the reacquisition process.

6. AIDED TRACKING

Useful GPS data is obtained only while the carrier-tracking and/or code-tracking loops are locked onto the desired signals. When any tracking error becomes too large, the correlation detector becomes excessively nonlinear and its effective gain is accordingly lowered. Progressive increases in the tracking error and attendant reductions in the detector gain lead to a complete loss of lock and to a complete loss of that component of the GPS data.

Loss of lock can occur because of excessive tracking error in response to radio noise, perhaps due to jamming. Lowering the loop bandwidth lowers the noise-induced tracking error, but also acts to increase errors in response to dynamics in the pseudorange variable that is to be tracked. These dynamics can be due to vehicle dynamics, to oscillator dynamics, and to perturbations in the propagation path. By utilizing the IMU to measure the short-term signal dynamics due to translation and rotation of the vehicle and by supplying the IMU data as an aiding signal to the tracking loop, the tracking-error components due to vehicle dynamics can be substantially reduced. With reduced errors due to vehicle dynamics, the bandwidth of the loop may be reduced to further attenuate jamming. Hence, IMU aiding of the tracking loops increases the antijamming margins of the tracking loops in dynamic environments[41-45].

Generally, aiding signals are introduced as shown in Figure 3, where τ is the time delay to be tracked, $\hat{\tau}$ is the estimated value of τ, and $\dot{\tau}_{aid}/K_{FC}$ is an aiding signal. Alternatively, with exactly the same effect on the detected error signal $D(\tau - \hat{\tau})$, the aiding signal τ_{aid} could be added to the feedback signal as shown by the dotted line in the figure. If τ_{aid} were subtracted from the input τ instead of being added to the feedback signal $\hat{\tau}$, its effect on the output of the detector would be the same. Hence the effect of aiding is to make the loop track an effective pseudorange signal $\tau_e \equiv (\tau - \tau_{aid})$. The IMU can supply an aiding signal τ_{aid} that matches the high-frequency components of τ quite well so that $\tau - \tau_{aid}$ contains smaller high-frequency components than τ. Then, with aiding, the bandwidth of the loop can be lowered in order to attenuate jamming noise while still maintaining linear operation of the detector.

In order to avoid loss of lock, the tracking error of each carrier-tracking loop should be kept well within the small value of 0.1 ft (phase-locked) or 0.05 ft (Costas). In order to provide this accuracy with narrow tracking-loop bandwidths during high-dynamic maneuvers of tactical aircraft, aiding velocity signals must be delivered from

the IMU to the carrier loops very frequently, on the order of every 10 ms, and transport delays must be accounted for. Because data are generally not available from the INS mechanization algorithms this frequently, the aiding data should be obtained directly from the IMU, or from its strapdown processor, as shown in Figure 2.

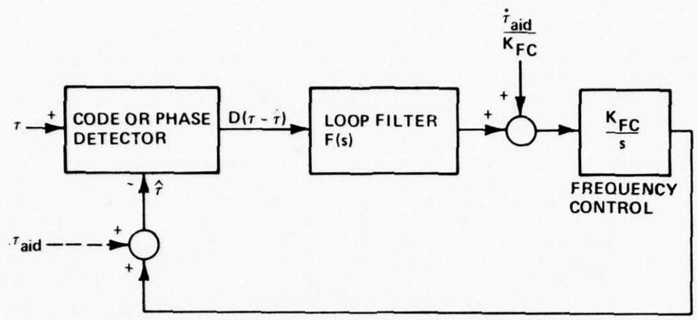


Figure 3. The use of an aiding signal in a tracking loop

Transport delays can occur in the implementation of the tracking loop as well as in the implementation of the aiding algorithms. The aiding signal for each loop is generally a series of increments $\Delta\tau_{aid}$, each of which, when implemented at some accurately specified time in the future as the frequency command to the loop over an interval Δt, represents the predicted change $\Delta\hat{\tau}$ in the signal to be tracked over that interval. Past values of acceleration and attitude data from the IMU are used as inputs to an extrapolation algorithm. The output of the extrapolation algorithm is the predicted $\Delta\hat{\tau}$, which is delivered to the tracking loop. The extrapolation algorithm can be very simple, but should be repeated very frequently, every 10 ms, or so for each carrier loop. Because the code-loop tracking errors can be on the order of tens of feet, much larger than the allowable carrier-loop errors, the problems of transport lags and sampling periods are relatively insignificant for code-loop aiding. Velocity information from the INS mechanization algorithms may be used for aiding the code loops.

The parameters for the coordinate transformation and scaling of the translational aiding data can be obtained from the integrated navigation filter, as shown in Figure 2. These parameters, relating to the bearings of lines of sight to the GPS satellites, and relating to the orientation of the gyro-stabilized accelerometer frame, change slowly and need to be updated only infrequently. With these parameters being known, the incremental, gyro-stabilized, accelerometer data are scaled and transformed into changes in velocities along the lines of sight. The transformation parameters for attitude data must be obtained from the IMU. Then incremental attitude data from the IMU are transformed into changes in ranges from the antenna(s) to the satellites. The transformations of data on translations and rotations are simple, but must be performed rapidly with proper time-tagging.

The transformation, scaling and extrapolation functions for inertially-aided GPS tracking require special software not needed for the unintegrated systems. These will be warranted in some applications because of the resulting improvement in antijam performance.

In high-dynamic maneuvers of tactical aircraft, the effects of acceleration on the frequency of the GPS crystal oscillator can be a significant cause of residual tracking error in inertially aided code or carrier tracking loops[42-47]. The frequency shifts appear as pseudorange-rate inputs to all the tracking loops and are not directly detectable by the IMU. Specifications for currently available crystal oscillators are no lower than 10^{-9}/g, which is equivalent to a pseudorange rate input of 1(ft/s)/g. For a code loop operating with a time constant of 30 seconds, a 2-g input to such an oscillator would lead to a 60-ft tracking error in addition to other dynamic error components. Hence the oscillator sensitivity limits antijamming performance of the code loops. It affects the antijamming performance of IMU-aided carrier loops too, especially during high-jerk maneuvers. Fortunately, it appears that crystal oscillator assemblies with sensitivities below 10^{-10}/g are on the horizon[48].

IMU errors also limit the antijam performance of aided code and carrier tracking loops. For the carrier loops, care must be taken to keep tracking errors due to attitude errors, bending modes, etc., to less than about 0.1 ft, with loop response times of about 0.5 s, or less, during severe attitude maneuvers. With one milliradian attitude errors, a lever arm length of much less than 100 ft would not be a problem, but structural flexure between the antenna and the IMU could be a problem. Either gimballed or strapdown IMUs are capable of supplying sufficiently accurate attitude data in most applications. For the gimballed IMUs the pacing requirement is on the accuracies and update rates of the gimbal angle encoders. References (44-45) provide quantitative data on the effects of a variety of error sources on the performance of IMU-aided carrier loops under jamming and with high dynamic maneuvers. Improvements of 10 dB or so in AJ performance can be achieved through inertial aiding of carrier loops.

For aiding the code loops, the ability of the INS to measure dynamic translational motions sufficiently accurately so that tracking errors are much less than 100 ft with loop response times on the order of 1-100 s is the dominant requirement. Only high-quality stable-platform IMUs are currently capable of successfully aiding code loops with time constants on the order of 100 s during high-dynamic tactical aircraft maneuvers. The code-loop time constant determines the ultimate antijam margin. Hence, the antijam performance of the integrated GPS inertial system will be strongly dependent upon the accuracy of the IMU over the full range of accuracy of available designs. This conclusion is true to an even greater extent when navigation performance after loss of lock, and reacquisition performance under moderate jamming, are considered.

Aided tracking loops provide GPS data that are corrupted to some extent by the aiding signals. This corruption has been reported in the literature[(49)] as sometimes leading to problems of instabilities in navigation algorithms designed to receive pure GPS data, and some solutions have been proposed. Because the aiding signal τ is effectively high-pass filtered by the tracking loop, subtracting a similarly high-pass filtered version of the aiding signal from the output of the aided tracking loop is an alternative approach to this problem. Some extra hardware or software is required for creating the high-pass-filter function.

Aided code loops will have very long time constants in order to maximize anti-jamming performance. When the time constants are longer than the sampling period of the navigation filter, the data samples will be correlated. Prewhitening filters have been implemented in some designs to reduce the correlations and thereby to improve the performance of the Kalman navigation filter.

7. ADAPTIVE TRACKING

When the conditions of signal dynamics and jamming noise are such that maintaining lock is difficult, choosing the tracking-loop parameters to optimize performance is important. Because the conditions usually cannot be predicted in advance, the parameter adjustment is best done adaptively. The availability of data from an IMU greatly facilitates the adaptation process, as indicated in Figure 2.

The best values for the parameters of each loop are dependent upon the dynamic model for the effective pseudorange $\tau_e \equiv (\tau - \tau_{aid})$ being estimated (tracked) by the loop, the signal-to-noise ratio, and the covariances of the tracking loop's estimates of the model states. In an integrated GPS-inertial system, data from the INS can be used directly to determine the time-variable parameters in a model of the effective pseudorange dynamics. Data from the INS can also be used, together with data from the receiver to determine the signal-to-noise ratio. From this information, together with initial values for covariances of the tracking error states, the succeeding tracking-error covariances and optimum tracking loop parameters can be calculated.

Each optimally adaptive, n^{th} order, tracking loop can be considered as a Kalman filter operating on noisy measurements of τ_e, which is generated by an n^{th} order dynamic process model with white noise sources. Once the model is characterized adaptively, the optimal tracking-loop parameters can be obtained from the standard Kalman algorithms. For example, for a third-order carrier tracking loop, an adequate process model might be $\dot{\tau}_e = v_e$, $\dot{v}_e = a_e$, $\dot{a}_e = -a_e/T + n_j$, where n_j is a white noise source with power spectral density N_j, and T is the correlation time of the acceleration state. The variations in effective pseudorange dynamics could be represented by variations in N_j.

If the loop is unaided, this process represents the full pseudorange dynamics. Then accelerometer and attitude outputs from the IMU can be used to designate appropriate values of N_j. If the loop is aided by the INS the process represents the dynamics of aiding error. Then the accelerometer and attitude outputs are used with the INS and oscillator error models to obtain appropriate values of N_j. For example, if the uncertainties in alignments of the stabilized axes are 10 milliradians rms, 1 percent of the acceleration indicated by the IMU could be allocated to the effective pseudorange dynamics. The uncertainties of the INS parameters will vary only slowly (as solutions of the navigation filter), but the accelerometer and attitude data can vary rapidly, and must be converted into selections of values for N_j very rapidly so that the loop will be able to adapt itself in time to follow high-jerk dynamics. The algorithms for selecting values for N_j must be very simple to be practical. The values of N_j should be appropriately bounded to prevent selection of excessively large or small tracking-loop bandwidths, but need to be only approximate indications of levels of dynamics within these bounds. For a first-order aided code-tracking loop, a simple first-order model for the effective pseudorange dynamics would be chosen.

The measurement of the effective pseudorange τ_e by the code or carrier detector is corrupted by noise n_τ. The power spectral density N_τ of n also is needed for deriving the optimum tracking-loop parameters. The value of N_τ is proportional to the spectral density of the radio noise (perhaps due to jamming) divided by the signal power. The radio noise density can be measured directly and rapidly by the receiver. The signal power is a weak function of the angle of the satellite above the horizon and

CIVIL APPLICATIONS OF NA STAR GPS

Glen A. Gilbert*
Glen A. Gilbert & Associates

a strong function of the attitude of the aircraft. By utilizing attitude data from the IMU and the known antenna patterns of the vehicle, the signal power can be predicted. Since the signals will usually be at full power, it may be sufficient to characterize each of them merely as present or absent. Because malfunctions or intervening fixed structures, such as mountains, can cause unpredicted blackouts of some signals, the receiver must also estimate the signal power that is actually received from each satellite. But this estimate requires code demodulation and, hence, is not valid when the code detector range is exceeded. It also depends upon the parameters of the measurement filter, which should themselves be optimized adaptively. Although the response time may be slow, provision should be made for the signal power measurement by the receiver to provide a "signal-absent" override to the adaptive algorithm, and to satellite-selection and navigation algorithms. The values for N should be bounded so that unreasonably large or small values are not utilized.

With the parameters N_j (representing effective pseudorange dynamics) and N (representing measurement noise) determined, Kalman solutions for the gains in the tracking loops can be obtained. The solutions are straightforward, but are time consuming when performed rapidly enough to be accurate for a loop with wide bandwidth. Time-scales approximation techniques have been applied to the simplification of a numerical solution of the Kalman equations for a second-order tracking loop(50), thereby providing a practical approach to that part of the real-time adaptive tracking task.

The solution for optimum tracking loop parameters has been extended to cover the effects of code-detector nonlinearity and to cover the possibility of also varying the code-detection range optimally so that tracking and acquisition can be handled with improved performance by a single adaptive tracking process.(51).

Improvements in antijam performance can also be obtained by adaptively varying the predetection bandwidth. The predetection bandwidth of each Costas carrier-tracking and incoherent code-tracking loop can be varied in proportion to the uncertainty in the effective pseudorange rate. Each bandwidth should be large enough to pass the carrier and data (if present), but otherwise should be as small as possible in order to minimize signal suppression when the predetection signal-to-noise ratio is less than unity. The improvement in performance through adaptation is most pronounced during acquisition, or when the code loop is data-aided so that its predetection bandwidth can be small(42). If the carrier loop is in lock, the predetection bandwidths of the carrier loop and its associated code loop should be proportional to the frequency uncertainty of carrier-loop tracking, as indicated by the Kalman adaptive tracking equations for the carrier loop. When the carrier loop is not in lock, and both the carrier and code loops are IMU-aided, the much slower navigation filter solutions for uncertainties in clock frequency and vehicle velocity can be utilized.

It is clear that the availability of data from an IMU greatly facilitates the process of adapting the GPS receiver to different operating conditions. An approach to the process has been suggested above. However, further exploratory development work is needed to develop complete sets of adaptation algorithms of different levels of complexity and to determine their performance advantages in specific applications.

8. ANTENNA CONTROL

In order to enhance the antijam performance of the integrated GPS inertial system, beam-pointing antennas can be employed(38). The beams must be pointed at the GPS satellites throughout aircraft maneuvers. Control for the beam pointing is generated on the basis of data from the IMU on the attitude of the vehicle and data from the GPS navigation processor on the position of the vehicle with respect to the satellites. The attitude data must be updated rapidly enough to allow predictions by the antenna control processor of the attitude of the vehicle with errors substantially less than the beamwidth during all attitude maneuvers. The position data, used to calculate the bearings to the satellites, need be updated only rarely.

9. CONCLUSION

Benefits and means of integrating GPS and inertial systems have been described. Emphasis has been on the data to be transferred and the operations to be performed in attaining varying degrees of integration. The intent has been to provide an understanding of the mechanisms and degrees of complexities involved, and to provide perspective on the technical issues that are involved in the integration problem. References have been supplied to the wealth of detailed material available in the literature on this topic.

It is clear from the discussion that very substantial performance improvements can be obtained through integration of GPS and inertial systems in comparison to what can be achieved by either system alone. Some of the improvements are only achieved through substantial increases in system complexity.

ACKNOWLEDGEMENT

The author is grateful for the advice and consultation of Dr. Bernard Kriegsman, Dr. James Negro, Mr. W. Michael Bowles, and Mr. William Stonestreet during the preparation of this paper.

interrupted, but again the implications would be large for the possible returns. Discontinuance or denial of NAVSTAR GPS precision signals would also have wide implications. Presumably the Department of Defense would want to deny the precise information in specific geographical areas although a large volume would necessarily be affected. It is conceivable that precise information could be denied Europe, Africa, or Asia without denying the signal to the United States by tran

1. W.J. Klepcznyski, ed., NAVIGATION, Vol. 25, No. 2, Summer 1978.

2. Federal Radio Navigation System Plan, Preliminary Draft, Office of Telecommunications Policy, Fall 1977; to be published after coordination.

3. Final Report on GPS-JTIDS-INS Integration Study, to be issued in Summer 1978 by The Charles Stark Draper Laboratory under contract F33657-77-C-0672.

4. R.Q. Lee and G.R. England, "The Digital Airplane", Astronautics & Aeronautics, January 1978, pp 58-64.

5. C.S. Draper, W. Wrigley, J. Hovorka, Inertial Guidance, Pergamon Press, N.Y., 1962.

6. G.R. Pitman, ed., Inertial Guidance, John Wiley & Sons, Inc., N.Y., 1962.

7. C.F. O'Donnell, ed., Inertial Navigation Analysis and Design, McGraw Hill Book Co., 1964.

8. C. Broxmeyer, Inertial Navigation Systems, McGraw Hill Book Co., N.Y., 1962.

9. M. Kayton, W. Fried, Avionics Navigation Systems, John Wiley & Sons, Inc., N.Y., 1969.

10. W. Wrigley, W.M. Hollister, W.G. Denhard, Gyroscopic Theory, Design, and Instrumentation, MIT Press, Cambridge, MA, 1969.

11. J. Sciegienny, R. Nurse, J. Wexler, P. Kampion, "Inertial Navigation System Standardized Software Development, Final Technical Report", Charles Stark Draper Laboratory, Report #R-977, Vols. 1-4, June 1976.

12. K. Britting, Inertial Navigation Systems Analysis, Wiley-Interscience, 1971.

13. J. L. Farrell, Integrated Aircraft Navigation, Academic Press, N.Y., 1976.

14. G.T. Schmidt, ed., AGARD Lecture Series #95, Strapdown Inertial Systems, Theory and Application, NATO AGARD Report LS-95, June 1978.

15. C.T. Leondes, ed., Theory and Applications of Kalman Filters, AGARDOGRAPH No. 139, NATO AGARD, Feb. 1970, AD704306.

16. Bryson, A.E., and Ho, Yu-Chi, "Applied Optimal Control", Blaisdell Publishing Co., Waltham, MA, 1969.

17. G.T. Schmidt, ed., AGARD Lecture Series #82 on Practical Aspects of Kalman Filtering Implementation, North Atlantic Treaty Organization, Advisory Group for Aerospace Research and Development, Report AGARD-LS-82, March 1976, ADA024377.

18. A. Gelb, ed., Applied Optimal Estimation, MIT Press, Cambridge, MA, 1974.

19. C.T. Leondes, ed., Filtering and Stochastic Control in Dynamic Systems, Vol. 12 of Control and Dynamic Systems, Academic Press, N.Y., 1976.

20. Kaminski, P.G. and Bryson, A.E., "Discrete Square Root Filtering", IEEE Transactions on Automatic Control, Vol. AC-16, No. 6, 1971.

21. Carlson, N.A., "Fast Triangular Formulation of the Square Root Filter", AIAA Journal, Vol. 14, No. 9, 1973.

22. Bierman, G.J., "Factorization Methods for Discrete Sequential Estimation", Academic Press, N.Y., 1977.

23. Hybrid Navigation Systems, AGARD Conference Proceedings, No. 54, Jan. 1970, US Government Report No AD-701775, pp. 14-1 thru 14-20.

24. D.E. Gustafson and B.A. Kriegsman, "A Guidance and Navigation System for Automatic Station Keeping in a Earth Orbit" Journal of Spacecraft and Rockets, Vol. 10, No. 6, June 1973, pp 369-376, Draper Laboratory Report E-2534, August 1970.

25. B.A. Kriegsman and Q. Yee, "Shuttle Navigation System for Entry and Landing Mission Phases", Journal of Spacecraft and Rockets, Vol. 12, No. 4, April 1975, pp. 213-219.

26. B.A. Kriegsman, "NAVSTAR-Aided Inertial Navigation Studies", Draper Laboratory Report R-1115, Sept. 1977.

27. C.A. Wolfe, "NAVSTAR/GPS Navigation Analysis and Algorithm Development Study", Final Report OC-R-76-0564-1, ORINCON Corp., AD-A-034 873, 30 November 1976.

28. R.A. Clark, M.A. Needler, "Performance and Adjoint Sensitivity Analysis of the NAVSTAR Global Positioning System", NAECON '76 Record, pp. 834-840.

29. K.A. Myers, R.R. Butler, "Simulation Results for an Integrated GPS/Inertial Aircraft Navigation System", NAECON '76 Record, pp. 841-848.

with code-division multiple access to each individual satellite, precision ranging measurement on the code modulation in continuous real time, significant processing gain against signal interference, and inherent multipath discrimination.

30. A.J. Brockstein, "GPS Kalman Augmented Inertial Navigation System Performance", NAECON '76 Record, pp. 864-871.

31. D.W. Candy, "A NAVSTAR/GPS Simulator", NAECON '77 Record, pp. 323-329.

32. B.A. Kriegsman, K.B. Maher, G. Prado, and G.W. Hurlbut, Deep Ocean Mining Site Navigation System Evaluation, C.S. Draper Laboratory Report R-1049, January 1975, presented at ION National Marine Meeting, Nov. 1977.

33. Damoulakis, J.N. and Hinderer, J., "A GPS-HDUE System Configuration with Increased Processing Power on Range and Range-Rate Measurements", Position Location and Navigation Symposium, Institute of Electrical and Electronics Engineers (IEEE), 1-3 November 1976, San Diego, CA.

34. Dressler, R.M. and Ross, D.W., "A Simplified Algorithm for Suboptimal Non-Linear State Estimation", Automatica, Vol. 6, pp 477-480, 1970.

35. P.S. Noe, K.A. Myers, "A Position Fixing Algorithm for the Low-Cost GPS Receiver", IEEE Transactions on Aerospace and Electronic Systems, Correspondence, March 1976, pp 295-297.

36. W.H. Foy, "Position Location Solutions by Taylor Series Estimation, IEEE Transactions on Aerospace and Electronic Systems, Vol. AES-12, #2, March 1976, pp. 187-194.

37. Spread Spectrum Communications, AGARD Lecture Series No. 58, NATO AGARD, July 1973.

38. G.L. Bjornsen, W.M. Hutchinson, "GDM/GPS Receiver Hardware Implementation", NAECON '77 Record, pp 303-309.

39. W.M. Stonestreet, A Functional Description of the NAVSTAR GPS Receiver Model-X, Charles Stark Draper Laboratory Report #R-981, Vol. 1, Rev. 1, Feb. 1977.

40. GPS User Equipment (UE) Orientation Course, Magnavox Report MX-125-C-US-7701, Magnavox Government & Industrial Electronics Co., Advanced Products Div., Torrance, CA, 1977.

41. P.P Yeh, "A Treatise on the Antijamming Margin of an IMU/Computer-Aided Global Positioning Navigation System", Aerospace Corporation Report TOR-0076(6473-01)-1, 15 Oct. 1975.

42. L.L. Horowitz, J.R. Sklar, "ECM Vulnerability of the GPS User Receiver in a Tactical Environment", Lincoln Laboratory Report XR-1, 7 May 1976.

43. E.H. Martin, "Aiding GPS Navigation Functions", NAECON '76 Record, pp. 849-856.

44. B.A. Kriegsman, D.B. Cox, Jr., W.M. Stonestreet, "An IMU Aided GPS Receiver ", CSDL Report P-490, June 1977, Presented at the 33rd Annual Meeting of the Institute of Navigation, Costa Mesa, CA., 22-24 June 1977.

45. B.A. Kriegsman, W.M. Stonestreet, and D.B. Cox, Jr., "Functional Requirements of the Interface Between the NAVSTAR GPS Receiver Model X and the Advanced Inertial Reference Sphere", Draper Laboratory Report R-981 (Air Force Report SAMSO TR 77-120), Volume II, December 1977.

46. R. Maher, "Oscillator and Frequency Management Requirements for GPS User Equipment", Proceedings of 1976 Frequency Control Symposium, June 1976, pp 384-389.

47. J.M. Przyjemski, P.L. Konop, "Limitations on GPS Receiver Performance Imposed by Crystal Oscillator G-Sensitivity", CSDL Report P-432, March 1977, NAECON '77 Record, pp 319-322.

48. J.M. Przyjemski, "A Compensation Technique for Acceleration-Induced Frequency Changes in Crystal Oscillators", Draper Laboratory Report P-606, NAECON '78 Record, May 1978.

49. R.W. Carrol, W.A. Mickelson, "Velocity Aiding of Non-Coherent GPS Receiver", NAECON '77 Record, pp. 311-318.

50. R.V. Ramnath, W.M. Bowles, "Asymptotic Analysis of a Class of Time-Varying Linear Filters", CSDL Report R-994, 25 Aug. 1976.

51. W.M. Bowles, "Optimal Acquisition and Tracking of a Pseudo-Random Code", Draper Laboratory Internal Memo 15L-76-005, 8 Jan 1976.

2.9 Navigation Accuracy

...performance envisioned for the Global Positioning System is a di-

APPLICATION OF GPS TO LOW COST TACTICAL WEAPONS

A. F. Schmitt*
J. J. Bowden
Teledyne Systems Company
Northridge, California 91324

SUMMARY

A navigation receiver employing the Global Positioning System (GPS) navigation satellites' signals is described which capitalizes upon initialization at the launch point so as to achieve considerable simplification and hence low unit cost. An excellent application of such a "mother-daughter" design is to the guidance of tactical weapons wherein the launch facility (ground launcher or aircraft) is presumed equipped with a full-capability GPS navigation set suitable for the initialization of the simplified receiver.

Once initialized with GPS data (satellite orbital parameters and clock corrections) and synchronized to track the received signals, this receiver is thereafter fully capable of autonomous navigation including signal reacquisition following any signal outage periods. In this design GPS data are not received from the satellites' transmissions and hence are not updated. Consequently mission time following launch must be limited to no more than an hour or so to minimize degradations in accuracy - a constraint of little concern in most tactical applications.

1.0 INTRODUCTION

1.1 GPS Application to Navigation

As a space-based radio navigation system NAVSTAR GPS is an attractive aid for the guidance of tactical weapons. Weapons using GPS guidance will provide all-weather, day-night operation and will be capable of world-wide, rapid response operation.

To be feasible however, a tactical GPS system must be inexpensive. One approach to achieve reduced costs of an expendable receiver and associated guidance equipment would employ the GPS signals in a retransmission mode for precise tracking of the tactical weapon to its target.[1] However, such a mode of operation sacrifices some of the major advantages which GPS guidance can offer. A tactical GPS guidance system should be autonomous (allowing launch-and-leave operation) and should serve an unlimited number of users simultaneously. Finally, GPS guidance is passive; it should not contribute to the detection or tracking of the weapon by the defensive forces. The retransmission mode, with its necessary tracking and command links, obtains none of these attributes.

In consequence, the retransmission mode is rejected as a means to low cost and alternative methods must be used to retain the principal virtues of GPS in a low cost receiver.

A GPS receiver navigates by making direct ranging measurements from its position to four of the eight-or-so transmitting satellites in view at any time. This quadrilateration is necessary because precise GPS time, as well as the three dimensions of user position, are unknown. The ranging modulation which makes this range measurement possible is a binary pseudo-random-noise code applied to the satellites' L-band transmissions. To navigate then, the receiver need only perform this ranging measurement and know precisely where the satellites are.

The satellites' precise positions are furnished to the user (the receiver) by data contained in a Navigation Message superposed upon each satellite's transmissions in another modulation far removed in frequency from the ranging modulation. With this second set of data, and a reasonably capable general purpose computer to complement the tracking circuitry, the navigation problem may be solved to present accurate three-dimensional positioning information to the user.

When one analyzes the above summary, a few complexities are found which require a typical GPS receiver to be capable of additional specialized functions. Provision must be made for initially searching for and acquiring the satellites' signals. If the highest precision is desired, the GPS receiver must directly measure systematic errors in L-band signal propagation. Furthermore, the receiver's computer will require a capability to select the optimum four satellites (of a 24-satellite system) to navigate from at any given time. Finally, it must be able to process the data of the

* This work was sponsored by the Air Force Armament Laboratory, Eglin Air Force Base, Florida, under contract F08635-76-C-0343.

to record information or dispatch vehicles. For command and control application, the location information is fed into a computer which makes decisions based on the input location information. First, let's discuss the location only uses of position information. In the case of fire departments, location information can be expected to be only of occasional use. Generally, fire equipment remains at the

Navigation Message, test for and correct errors, and update its own files when these messages are updated - as they are each hour. These features would all be present in what might be described as a "full-up" GPS receiver/navigator.

1.2 Mother-Daughter Approach to a Low Cost GPS Receiver

Achievement of the desired tactical GPS receiver cost and performance goals is expedited by capitalizing upon capabilities which will exist at the launcher or within the launch aircraft. These capabilities are used to effect certain prelaunch functions so that the tactical receiver is considerably simplified as compared to a full-up receiver. In addition, the receiver is rate-aided by an inertial guidance subsystem integral to the tactical weapon. Signal tracking performance is thereby enhanced under dynamic conditions (Figure 1).

The Mother-Daughter handoff of signal tracking deletes the necessity for the Tactical Receiver to carry a "Clear-Acquisition" PN code generator otherwise needed to acquire and track the 1 MHz C/A code. Instead, only the 10 MHz P-Code, used to achieve the precise satellite ranging measurements, must be generated in the receiver.

In the prelaunch period a "Mother-Daughter" philosophy of design calls for the launch facility or launch aircraft (the "Mother") to acquire and track the desired GPS satellite constellation. The Mother hands off the GPS signal synchronization, target coordinates and all needed GPS system parameters to the Daughter (the tactical GPS receiver) just prior to launch. This approach permits a number of dramatic simplifications in the tactical receiver. For example, the tactical receiver need not carry the capability to choose and acquire the GPS satellites, and it need not extract the Navigation Message which is superposed on each satellite's transmission. Eliminating the requirements to carry the GPS almanac, to solve satellite selection routines and to process Navigation Message data permits a great reduction in receiver computational capacity. Also, a coherent carrier tracker (Costas phase tracker) is not needed for extracting the 50 Hz Navigation Message from the satellites' signals.

Other simplifications achieved in the tactical receiver will be described later following a review of the design background for GPS receivers.

2.0 TACTICAL GPS RECEIVER DESIGN BACKGROUND

2.1 GPS Signal Structure

The GPS signal is basically derived from a very stable CW source to which two bit streams are applied by phase modulation. The bit streams are in phase quadrature, and each is a biphase signal. The modulation will be described in more detail later in this article after the following digression.

Consider a geo-stationary satellite in synchronous earth orbit transmitting an L band (now D band) fixed-frequency continuous wave signal stable, say, to 1 part in 10^{12}. An earth-based receiver should have no difficulty in receiving and phase locking to this signal assuming a reasonable link power budget has been provided. Since there are virtually no signal dynamics other than ionospheric and tropospheric aberrations, and no modulation, the receiver can be narrow banded with limits imposed principally by the phase noise(s) generated in the transmitter and the receiver.

Now consider the satellite in a circular polar orbit at half-synchronous altitude, such that its orbital period is 12 hours. Locate the receiver at the equator such that the satellite passes directly overhead (zenith). When the satellite rises into the field of view, it will have a component of velocity toward the receiver (positive). As it passes overhead this velocity will go through zero, and, as it recedes, the velocity component will become increasingly negative until the satellite sets. The apparent frequency at the receiver will differ from that which the satellite is transmitting owing to the doppler frequency shift, given approximately by the expression:

$$f_R = f_t \left(1 \pm \frac{V_{LOS}}{C}\right) \tag{1}$$

where:

f_R = received frequency

f_t = transmitted frequency

V_{LOS} = relative velocity along the line-of-sight between transmitter and receiver

C = velocity of propagation of radio waves

21-6

... indicate a hopeful trend. Additionally, the bus command and control situation can be auto-

Still, an earth-based receiver will have no difficulty in phase locking to this signal, but as yet, the only information derived is that the satellite is in orbit and is transmitting. If a precision frequency source is introduced into the receiver comparable with that in the satellite, the first genuinely useful piece of information can be obtained, viz., the relative velocity between the transmitter and the receiver. This is the beginning of a navigation solution. In 1977, the first satellite transmitting GPS signals, the NTS-2, yielded the plot of the measured doppler frequency versus time seen in Figure 2. The data differ somewhat from the example previously cited, in that a) the NTS-2 is in an orbit inclined 63° with respect to the equator (rises in the northwest and sets in the southeast), and b) the tracking station was actually located 34° north of the equator. These factors produce the observed asymmetry in the doppler plot. Note that this orbital inclination increases the exposure (useful tracking time) of a given satellite to a fixed earth-based observer relative to a polar orbit.

The tracking receiver for the above results was designed to make a direct P code acquisition of the GPS signals from the NTS-2 satellite and to remove the synchronous biphase P code modulation (ranging modulation) to permit carrier tracking. Note that while the P code modulation permits range measurement, the doppler shift of the carrier provides relative velocity data.

One should not dismiss the task of direct P code acquisition and removal of the biphase modulation too lightly. In order to perform these functions it is necessary to synchronize the receiver-generated P code phase and phase rate with the received signal, and thereby achieve signal correlation. The code phase error must be less than 100 nanoseconds in this seven day long code segment. When correlation has been achieved, and the data referred to Universal Coordinated Time (UTC) a direct range measurement to the satellite has been made. Figure 3 is a plot of range to the NTS-2 satellite measured at the same tracking station as was the doppler plot of Figure 2 (assuming no satellite clock offset from UTC).

2.2 The Codes

The L_1 carrier sustains two pseudorandom codes: the clear/acquisition or C/A code and the precision or P code, as well as a navigation message.

Each code is a bit-stream* of ones and zeros, synchronously modulated onto the carrier such that a carrier phase reversal occurs at each transition from a one bit to a zero bit and conversely from a zero bit to a one bit. The two codes are modulated onto the carrier in phase quadrature.

Since the codes are pseudorandom, each contains groups of two ones, three ones, four ones, etc. and like groups of two zeros, three zeros, four zeros, etc; hence there are periods within the code where the carrier phase remains constant over several chips.

The C/A code is a Gold code of 1023 chips length, clocked at 1.023 MHz. Thus its epoch time is one millisecond. Note that dividing the carrier L_1 by the C/A clock rate reveals 1540 carrier cycles per C/A chip. A receiver is thus assured of 1540 carrier cycles between successive phase reversals. The C/A code is of little interest to the tactical missile receiver, since the primary purpose of this code is to facilitate initial signal acquisition. The tactical user is privy to initialization data from a receiver which is already tracking the more precise P code.

2.3 P Code - The Time Domain

The P code is a very long code (267 days) which has been divided into 37 segments. Twenty-four of these are assigned to the GPS satellites. Each satellite transmits a 7 day segment which is reset to its initial state weekly at midnight Saturday UTC. The chipping rate (clock rate) of this code is 10.23 MHz, which produces only 154 carrier cycles in each chip. Figure 6 is a representation of the L_1 *P signal structure.

One may envision the carrier as a sinusoidal wave form with 180° phase reversals at P chip transitions. The period of each P chip is $\frac{1}{10.23}$ μsec or ≈ 98 nanoseconds as transmitted. Typically, the period of the received signal will be a slightly longer or shorter than that transmitted, owing to the doppler effect. The receiver must produce a replica of the code to be tracked which is correct both in phase, as measured from midnight Saturday, and in period. This calls for a code generator which can be rapidly slewed to the proper phase and a variable clock drive (sometimes called VCC**) which can be controlled to match the received code chip period. One may also view the change in code chip period as an apparent change in the received code chipping rate or code frequency. Eq. 1 applies, using the transmitted code chipping rate of 10.23 MHz. Since it is known that there are 154 carrier cycles in each chip, the code doppler is $\frac{1}{154}$ of the L_1 carrier doppler.

* Code bits are generally referred to as chips.

** VCC = Voltage controlled clock.

The L_1 carrier frequency shift can be seen on Figure 3 to exceed 4,000 Hz to an earth-fixed observer. This corresponds to a relative line-of-sight velocity of about 800 m/s arising from combined satellite and earth motion. A receiver mounted in a typical jet aircraft can readily sustain an additional 800 m/s velocity. Depending on whether the aircraft is approaching the SV or receding from it, the net LOS velocity could produce a doppler shift as high as $\pm$8000 Hz. Here one begins to see the dynamic aspects of the tracking requirement. Obviously, both carrier and code can sustain rapid dynamic changes. For the tactical mission, the trick is to be able to handle such signal dynamics under heavy jamming conditions.

2.4 P Code - The Frequency Domain

It is of interest to examine the signal spectra in the frequency domain before addressing the subject of jamming. Figure 4 is a series of spectrum analyzer photos which show the $\frac{\sin x}{x}$ distribution of signal energy produced by the pseudorandom biphase P code. The C/A code is omitted for clarity of presentation.

After correlation, that is, mixing of the received signal with the receiver-generated replica of correct phase, an IF signal is developed, as shown in Figure 5.

The process of correlation essentially collapses the spread-spectrum to nearly a line spectrum. If the navigation message data bits have not been removed from the IF signal, the "line" spectrum will be less than 100 Hz wide. A nominal spectrum width of 60 Hz is estimated, based on the assumption that the data bits are random, and the Fourier transform of the data bit sequence is also a $\frac{\sin x}{x}$ distribution.

2.5 Interference Resistance of GPS

The simplest form of interfering signal to produce (and to analyze) is a non-coherent continuous wave interference (CWI) line spectrum. Suppose that an in-band CWI signal is processed through the preselection filter, low noise RF amplifier and other stages which precede correlation together with the desired GPS signals, without limiting. At the point of signal correlation, the P code is removed from the carrier, usually in a double balanced mixer. Effectively, the received signal is multiplied by the receiver-generated P code replica and, when correlation is achieved, the demodulated carrier signal is recovered for further processing as previously shown. Note that the aforementioned non-coherent CWI signal is also multiplied by the P code replica. This action spreads the CWI line spectrum to a sin x/x distributed signal and it becomes a replica of the receiver-generated L. O. spectrum. Thus the power in the CWI signal is spread over 20 MHz. The post-correlation filter bandwidth is narrow relative to 20 MHz, since at most, it need only pass the carrier $\pm$ the carrier doppler. If $\pm$ 10 kHz bounds the carrier doppler extremes, then a 20 kHz bandwidth will suffice. Thus, only the very center of the now-spread interfering spectrum will be admitted to further processing stages, and it appears as noise to the receiver. In this step, a "processing gain" of 1000 (or 30 dB) has been achieved. If the received signal could ultimately be processed through a 1 Hz filter, a processing gain of 70 dB could be achieved, which would mean that the jamming signal is suppressed by 70 dB.

Figure 7 is useful in envisioning the overall signal levels, noise levels, and jamming levels that a receiver can expect. The effective receiver noise density N_{oe} is essentially the jamming signal density J_o when heavy jamming is present. J_o is the jamming power divided by the spreading width of the receiver-generated P code replica.

2.6 Code Tracking in a GPS Receiver

The P code for a given satellite is a 7 day long segment of a much longer pseudorandom sequence, clocked at 10.23 MHz rate.

Let the value of a "1" chip of the P code be +1, and the value of a "zero" chip be -1. If the base-band P code is then introduced into an integrator, the average value will approach zero, since the code contains almost an equal number of +1's and -1's. If this code is multiplied by itself (or, as in the case of a receiver, by a replica) without phase error, then a maximum average dc value will be observed. If the dc value of each chip is unity, then time integration of the product will produce a value equal to the number of chips in the integration interval. Figure 8 is a simple example based on a very short code. For the P code parameters integration of the product for 1 ms will produce a maximum number equal to $10.23 \times 10^6 \times 10^{-3} = 10,230$.

In the GPS system the navigation message bit clock rate of 50 bps (20 ms.) affects the choice of integration period. General purpose receivers must limit their initial integration periods to fractions of the 20 ms. data bit periods - say to 4 or 5 ms. periods. This arises from the uncertainties in the phase and the phase rate of the receiver-generated P code replica relative to the received signal, which exist at the time of initial signal acquisition. Once the signal has been acquired and is being tracked, these uncertainties disappear, and the integration period can be increased to the data bit period of 20 ms. If the sign of each upcoming data bit was known in advance, the bits could be added

to the receiver-generated replica P code by an exclusive OR process (as they are in the satellite) and the integration interval could be further increased, subject to limitations imposed by unpredicted receiver dynamics.

It is typical to insert a selected code phase error in the receiver-generated replica to permit code phase tracking. If the input signal is split into two channels and one channel is operated with the receiver code displaced 1/2 chip early, and the other channel 1/2 chip late, the, when the signal power in both channels is equal, the nominal code phase is correct. Note that the channel integrators will observe correlation for only 1/2 the integration period, since the P code is orthogonal to itself when it is 1 chip in error. This code tracking function can be performed as indicated with two channels operating in parallel, or sequentially, by alternately presenting a receiver-generated "early" code and then a "late" code to the correlator. The latter configuration is termed a tau-dither system.

2.7 Carrier Tracking in a GPS Receiver

Once correlation has been acnieved, and the code removed, carrier processing can be undertaken. Typically, in the more general-purpose GPS receivers, the carrier is introduced into a Costas loop to permit demodulation of the 50 bps navigation message. For a short range tactical application, the navigation message provides valuable data, but once the information has been obtained, the modulation produces undesirable band spreading of the carrier. This is especially the case for a short duration tactical mission. Without the message, a phase lock loop could be employed for carrier tracking, which offers up to 6 dB SNR advantage over the ordinary Costas loop. One may postulate a data-aided system wherein the navigation message is stored in the missile receiver and demodulated from the carrier along with the code - thus permitting the advantage of phase lock carrier tracking. Such a tracker would have to be capable of either Costas or phase lock modes, since in the event of a change of data in mid-mission the phase lock tracker could no longer sustain signal lock. An interesting alternative approach embodies a frequency tracker in lieu of the Costas tracker. Such devices have been designed as incoherent "energy trackers." Incoherent in this context means that signal tracking can be performed while a frequency error exists. If the output of such a device is taken to be proportional to the relative velocity along the line-of-sight from the receiver to the satellite, it will of course exhibit some error value. At the GPS L_1 carrier frequency a 1 Hz error is equivalent to 0.2 m/s velocity error. Such a system produces somewhat degraded accuracy with respect to Costas or phase lock loops, which sustain only phase errors, but no frequency errors.* The other side of the coin favoring the frequency tracker approach includes certain very significant advantages. Dynamics which can cause break lock in phase tracking loops cause only cycle slips in a frequency tracker. Such cycle slips contribute tracking error, but do not cause loss of signal tracking unless they are so frequent as to introduce a large frequency error. Hence, one may properly conclude that the frequency tracker has greater tolerance to dynamics for a given signal-to-noise or signal-to-jammer ratio.

The ultimate performance of the frequency tracker depends on the loop bandwidth. Inherently such devices can be designed to track with a signal-to-noise ratio of unity. If one postulates a 10 Hz bandwidth, then the minimum GPS signal-to-noise density that one can expect to track is 10 dB-Hz.

3.0 TACTICAL RECEIVER PERFORMANCE

3.1 Signal Dynamics and Tracking Loop Bandwidths

A receiver tracking a GPS satellite signal from a fixed location on the earth will experience limited signal dynamics as may be observed in Figure 2. The optimum loop bandwidths should be quite narrow owing to the limited rates of change of frequency and phase of the carrier and code. The narrower the loop bandwidth, the more jam resistant the receiver becomes. If the earth based receiver is installed in an aircraft capable of high dynamics, its optimum loop bandwidths will no longer be adequate for signal tracking. If the aircraft dynamics could be furnished to the receiver in advance, appropriate compensation could be provided, and narrow loop bandwidths could be maintained. One approach to predicting dynamics is to employ an inertial sensor to furnish velocity change data. With such inputs, the receiver needs only to have loop bandwidths wide enough to accommodate the residual errors in the inertially derived predictions, which will tend to approach the optimum loop bandwidths for the ground based receiver.

An important hardware savings is also available to the tactical receiver through the utilization of the inertial guidance subsystem (IGS) on-board the missile. Depending upon the quality of the IGS the tactical receiver hardware may be further reduced by limiting the number of receiver trackers. Fewer than four tracking channels can be time shared among the requisite four satellites to be tracked. This sequential tracking of satellites necessitates a reacquisition of a given satellite's signal during each sequencing period, and therefore presupposes rather accurate dead reckoning capability from the IGS of a maneuvering vehicle. Obviously, this design problem presents a classic trade study between numbers of receiver tracking loops (cost of electronics) and quality of the IGS (cost of inertial instruments).

* In Costas trackings systems uncompensated clock errors will in fact produce velocity errors when tracking data are converted to state vector representation.

As a final note one may observe that the combination of a GPS receiver with its precision position and velocity measuring ability with an inertial sensor whose basic measurements are accelerations, provides a truly synergistic relationship.

3.2 Tactical Receiver System Integration

Further hardware savings in the tactical receiver are achieved by relying upon the Mother receiver to make direct measurements of ionospheric propagation delay corrections. The Mother receiver will track satellite signals on two different L-band frequencies and perform a delay difference measurement to calibrate this effect. The results will be available to the Daughter so that the tactical receiver does not have to carry the hardware for receiving a second L-band channel. Figure 9 depicts the partitioning of Mother-Daughter receiver functions.

In spite of the simplifications discussed above, the tactical receiver still requires a significant amount of supporting computation and logical decision-making. It is necessary, therefore, to control the receiver with a general purpose computer. For this purpose any of several contemporary micro-computers are quite adequate. This computer will not only perform receiver control computations, but will also form the data interface between the receiver and the rest of the missile's avionics system.

From the receiver's viewpoint the data interface involves transferring position and velocity data in "GPS coordinates", i.e., ranges and range rates along the lines of sight (LOS) to the satellites. The computer obtains these measured data from the receiver and, in turn, rate-aids the receiver by sending to it velocity changes along each LOS as resolved from independent data supplied by the IGS. To use (and to supply) these GPS-oriented data the computer performs the GPS-to-geographic coordinate conversions. Such conversions require continuous, precise knowledge of the satellites' positions and other miscellaneous GPS data and hence the computer must perform the orbital computations for the satellite constellation being tracked.

3.3 Tactical Receiver Navigation Accuracies

The inherent navigation accuracies of the simplified tactical GPS receiver approach those of the GPS, i.e., an average 8 to 11 meter spherical error, worldwide, 24 hours-a-day. Table 1 shows the GPS error budget which leads to this performance figure (References 2, 3, 6). For a sequential tracking tactical receiver ranging accuracy is poorer and, presuming the missile flies up to 60 minutes with the same satellite constellation, the GDOP will change giving rise to greater error contributions.* Obviously, the extent of the latter error is dependent upon flight-time, the constellation initially chosen and even upon the direction of the flight. The atmospheric delay error will also grow with time of flight since after launch the propagation correction can only be modeled in the tactical receiver's computer.

The tactical receiver's ranging error measurement may be degraded, in a time-averaged statistical sense, by assumptions of an occasional loss of signal tracking, followed by a search and reacquisition. The Tactical GPS Guidance system does have a signal search and reacquisition capability permitting mission scenarios wherein the receiver is jammed for a time or the satellite LOS is obscured by maneuvers or terrain masking.

REFERENCES

1. Raab, F.H. et al., "An Application of the Global Positioning System to Search and Rescue and Remote Tracking," NAVIGATION: Journal of the Institute of Navigation, Vol. 24, No. 3, Fall 1977.

2. "Contract Definition Final Report for Control/User Segments, Vol. III, User Segment Description, Performance Error Budgets and RF Link Budgets," General Dynamics/Electronics Division, R-74-032, 28 February 1974.

3. Kishel, J. J., "Applications of Global Positioning System to Tactical Missile Guidance," R&D Associates, RDA-TR-4401-006, September 1975.

4. "System Segment Specification for the User System Segment of the NAVSTAR Global Positioning System, Phase I," U.S. Air Force Space and Missiles Systems Organization, SS-US-101B, 30 September 1974.

* Optimally, a constellation is chosen such that the GDOP is unconstrained at launch but is minimum at the target.

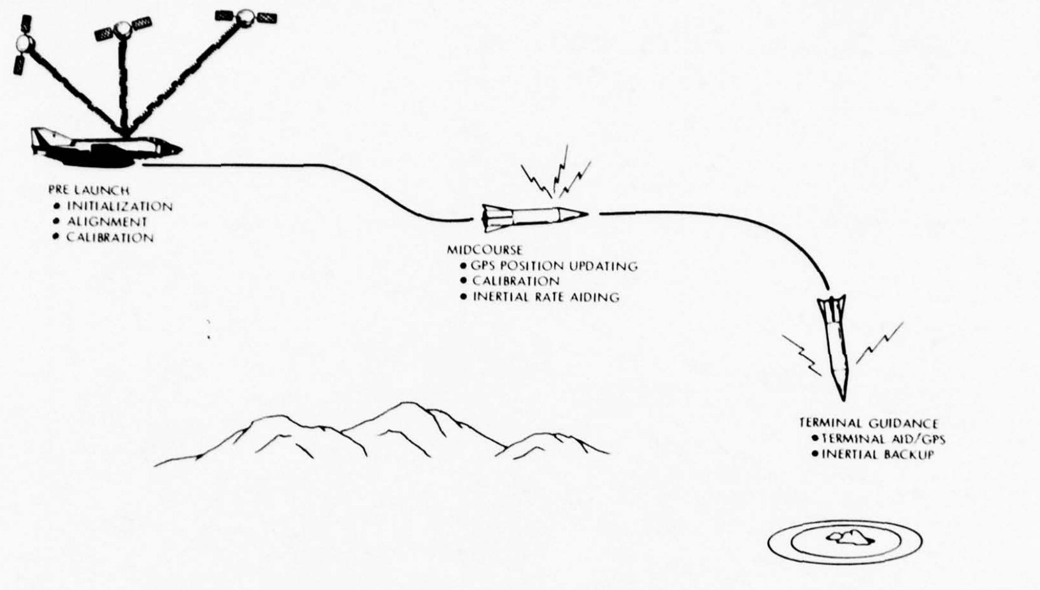


Figure 1. Tactical GPS Guidance Sequence

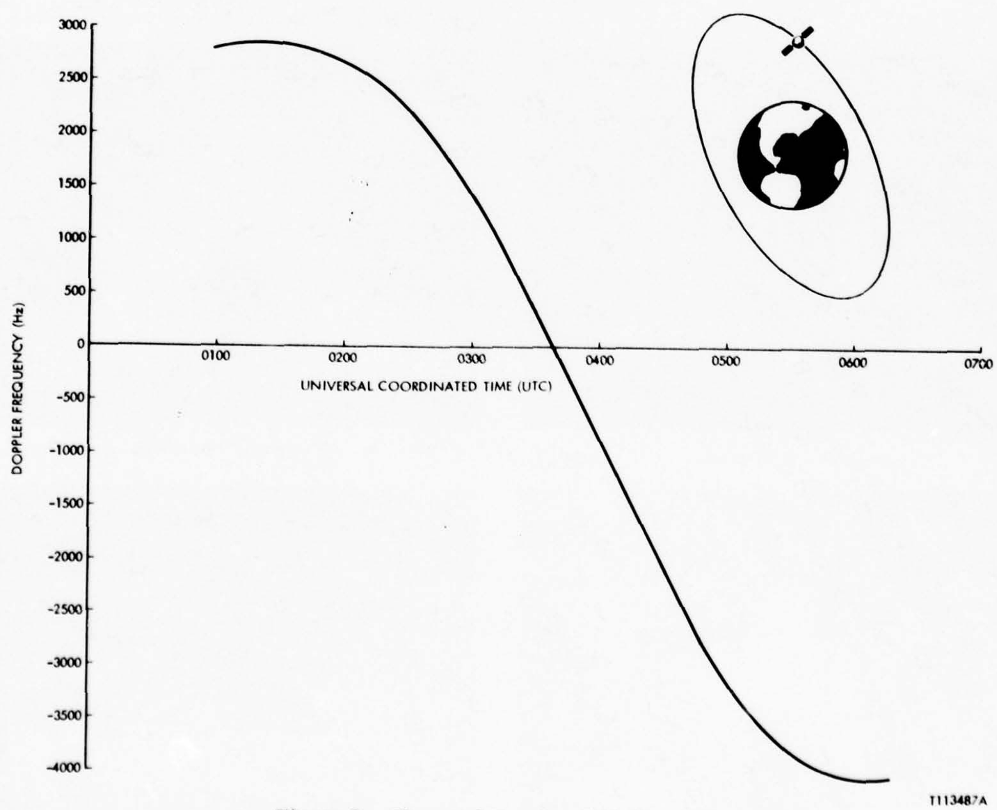


Figure 2. Observed Doppler Shift Time History for NTS-2. Ground Station at 34.2°N, 118.6°W

UNCERTAINTIES IN
SYSTEM PARAMETERS

Figure 2. Structure for integrated GPS inertial navigation system.

ERROR SOURCE	ERROR, METERS, 1σ
SATELLITE EPHEMERIS	1.5
UNCORRECTED ATMOSPHERIC DELAY	2.5
MULTIPATH	1.8
SATELLITES' UNMODELED CLOCK ERRORS AND SIGNAL DELAYS	.9
RECEIVER RANGING MEASUREMENT DEVIATION	1.5
R.S.S.	3.8
GDOP*	2.2 TO 3
ONE SIGMA SPHERICAL ERROR, METERS	8 TO 11

*THE POSITION GEOMETRIC DILUTION OF PRECISION - A FACTOR WHICH DESCRIBES THE DEGRADATION OF ACCURACY IN THREE DIMENSIONS DUE TO THE NON-OPTIMUM GEOMETRIC CONFIGURATION OF SATELLITES.

T116934

Table 1. GPS Navigation Errors for a Full-Capability Receiver

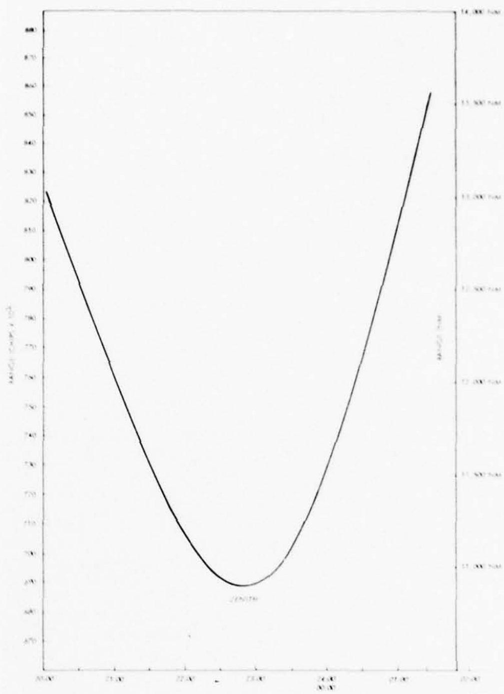


Figure 3. Observed Range Time History for NTS-2. Ground Station at 34.2°N, 118.6°W

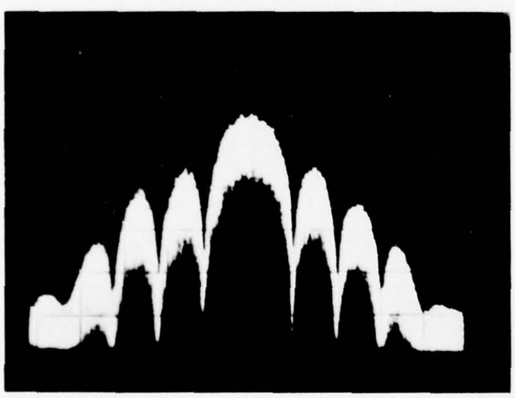


Figure 4a. GPS Signal L_1*P

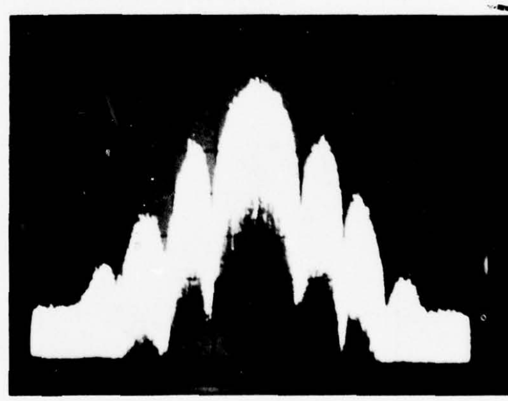


Figure 4c. Receiver-Generated L_1*P Replica

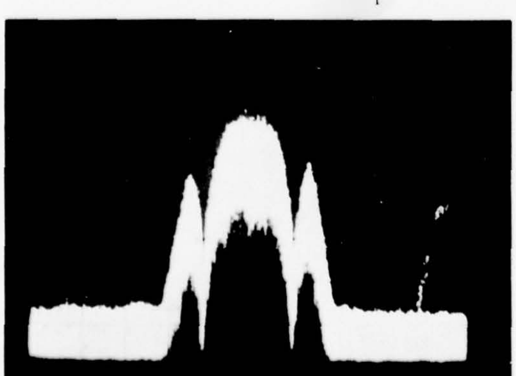


Figure 4b. GPS Signal L_1*P After Preselector Filtering

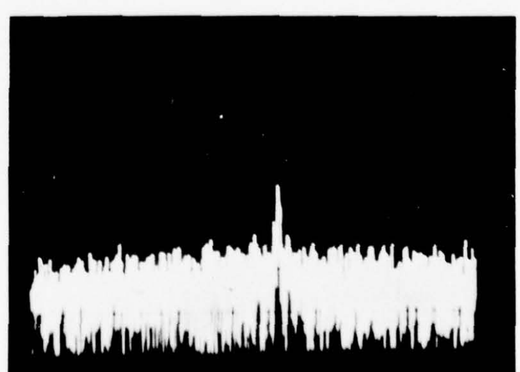


Figure 5. IF Signal After P Code Correlation

should be kept well within the small value of 0.1 Hz (phase-locked) or 0.05 Hz (Costas). In order to provide this accuracy with narrow tracking-loop bandwidths during high-dynamic maneuvers of tactical aircraft, aiding velocity signals must be delivered from

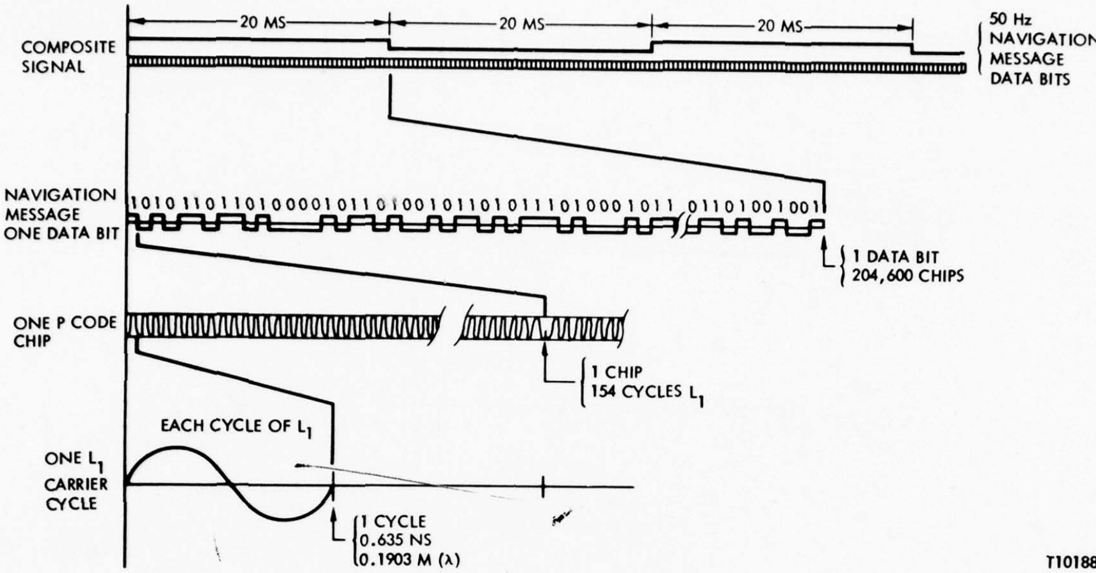


Figure 6. GPS - Signal Format - P Code

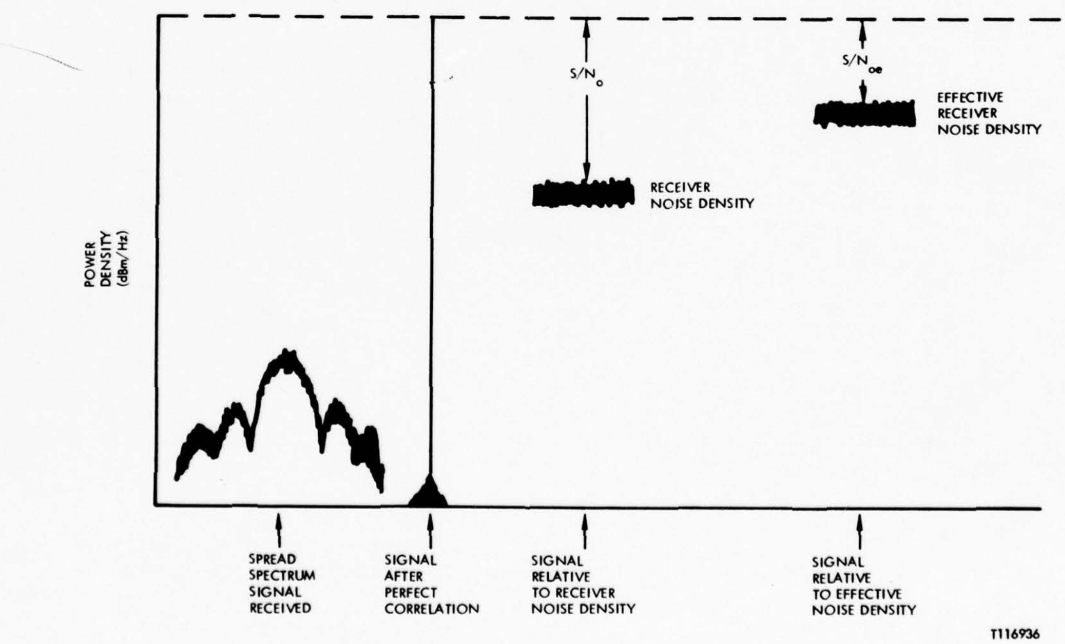


Figure 7. GPS Signal Power Levels

flexure between the antenna and the IMU could be a problem. Either gimballed or strapdown IMUs are capable of supplying sufficiently accurate attitude data in most applications. For the gimballed IMUs the pacing requirement is on the accuracies and update rates of the gimbal angle encoders. References (44-45) provide quantitative data on the effects of a variety of error sources on the performance of IMU-aided carrier loops under jamming and with high dynamic maneuvers. Improvements of 10 dB or so in AJ performance can be achieved through inertial aiding of carrier loops.

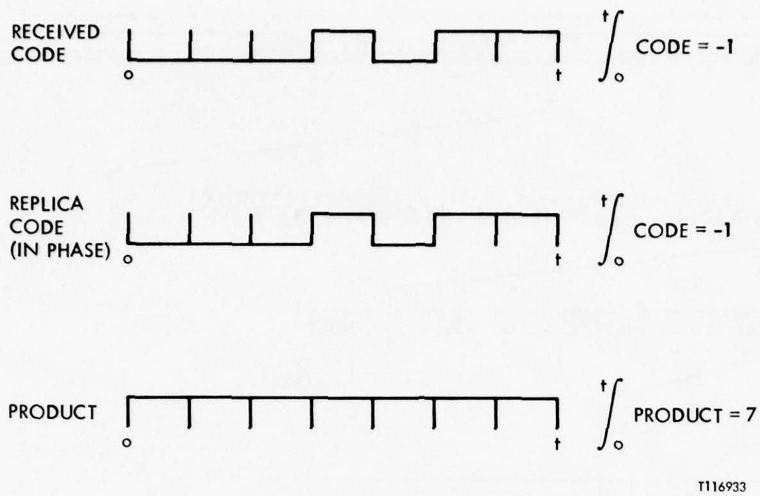


Figure 8. A 7 Chip Linear Sequence

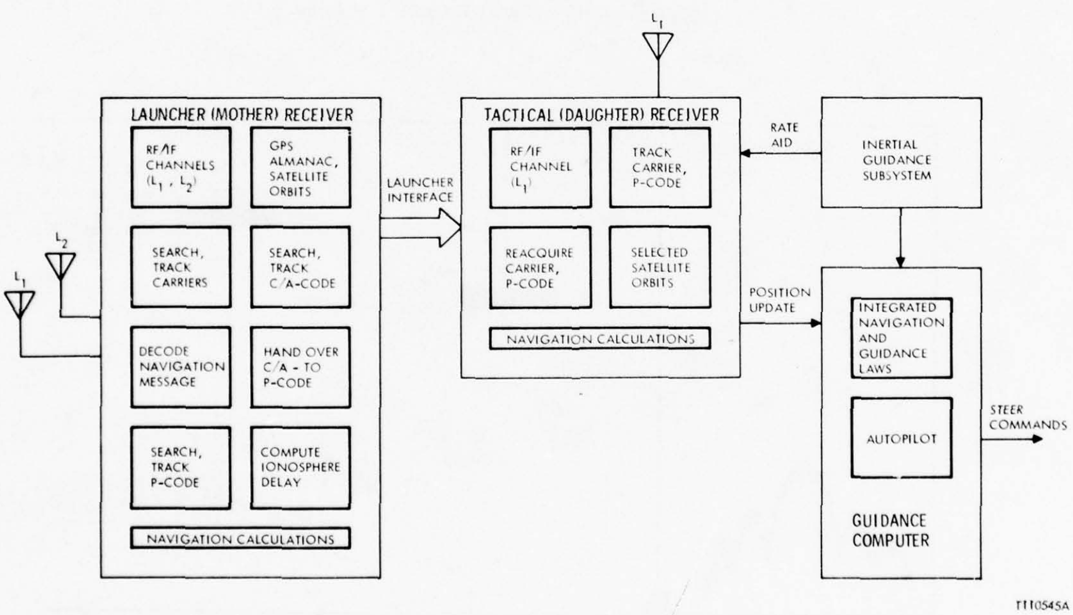


Figure 9. Tactical GPS Guidance Functional Diagram and Mother-Daughter Partitioning

is corrupted by noise n_T. The power spectral density N_T of n also is needed for deriving the optimum tracking-loop parameters. The value of N_T is proportional to the spectral density of the radio noise (perhaps due to jamming) divided by the signal power. The radio noise density can be measured directly and rapidly by the receiver. The signal power is a weak function of the angle of the satellite above the horizon and

CIVIL APPLICATIONS OF NAVSTAR GPS

Glen A. Gilbert *
Glen A. Gilbert & Associates
2500 Virginia Ave. N. W.
Washington, D. C. 20037

SUMMARY

This paper discusses various aspects of potential civil applications of the U. S. Department of Defense satellite based NAVSTAR Global Positioning System (GPS). Land, sea and air applications are covered. The conclusion is that the GPS offers many promising potential civil applications. However, certain questions will need to be resolved before international civil application of GPS can be expected to be implemented. Nevertheless, it is expected that these can be solved.

1. 0 INTRODUCTION

With respect to the availability by the U. S. Department of Defense of NAVSTAR GPS for civil applications, the following views and assumptions are exclusively those of the author. DOD will make available at all times to users in the civil sector information which will provide an accuracy in the order of 100 meters or less in three dimensions.

DOD is deploying NAVSTAR GPS as a positioning system in support of weapons delivery. The inherent value of the most accurate information - that is, in the order of 10 meters in three dimensions - is adequate for many military systems. Therefore DOD has stated that the military may deny the availability of the more precise signal under certain conditions. The signal would be otherwise available to the civil sector.

On the surface this policy appears to be forthright, specific, and unambiguous. However, there are a number of reasons to examine that policy in greater light as it pertains to a more or less guaranteed availability of the precise accuracy signal for civil use. First, in the interest of national security the military should have the capability to deny the use of such precise information to an enemy. In fact, under Executive Order 11161 the Department of Defense will assume control of the National Airspace System and the nation's navigation aids under certain emergency or hostile conditions. Therefore, DOD can in reality deny any navigation aid under the control and operation of the United States. Therefore, irrespective of a stated explicit policy for some portion of a specific system, under essentially the same conditions the military already may deny <u>any</u> information to civil users.

Policy is one matter, but what are the practicalities of availability and denial to civil users? There are at least two reasons why the coarse and precise signals will be available to civil users under normal, peace time conditions. First, it is likely that the Department of Defense will seek to involve the total national sector to improve the economy of the system to DOD. Second, as with the Navy Navigation Satellite System, any civil use of the system will stimulate an accumulating demand. The relative, good attributes of NAVSTAR GPS - precision, availability and economy - will become apparent to the civil and public sectors to include Congress so that there will be an ultimate requirement that DOD share the system in the fullest sense. It seems reasonable that sooner or later NAVSTAR GPS must become a national system or another satellite system would have to be deployed. The latter would undoubtedly be a relatively costly alternative.

What about under conditions of hostilities? DOD must be able and willing to carry out denial to make the threat of denial credible to an enemy. They should not be constrained to keep the signal on, and indeed they are not required to keep any navigational aid on when control of the nation's navigational aids is transferred to the Department of Defense. At the same time DOD must make the system available to its own forces, cooperating allied forces, and supporting reserve forces, and such civil carriers as the Contract Reserve Air Force. It does not seem practical that the navigation signal could be made available to such a wide number of friendly forces and denied to others except at great cost. Therefore, the Department can maintain a threat and capability to deny the system, provide the signal under almost any circumstance, and weigh national consequences at times when it believes it might be wise to deny the signal. Provided the civil sector agencies of the Federal Government join in (as they must) to advocate the system's use, there would obviously be only the direst circumstances when the precise signal (and only the precise signal) would be denied.

Even so, how would such denial be manifested? Individual VORTACs may be cut off independently (although it is not clear how such would be carried out practically). LORAN and OMEGA stations may be shut down easier, but with wider implications. TRANSIT satellites' transmissions presumably can be

---

* Co Authors:

E. H. Martin, Magnavox Government & Industrial Electronics Company

Denis Symes, Urban Mass Transportation Administration

Carl Mathews, Maritime Administration

substantial increases in system complexity.

ACKNOWLEDGEMENT

The author is grateful for the advice and consultation of Dr. Bernard Kriegsman, Dr. James Negro, Mr. W. Michael Bowles, and Mr. William Stonestreet during the preparation of this paper.

interrupted, but again the implications would be large for the possible returns. Discontinuance or denial of NAVSTAR GPS precision signals would also have wide implications. Presumably the Department of Defense would want to deny the precise information in specific geographical areas although a large volume would necessarily be affected. It is conceivable that precise information could be denied Europe, Africa, or Asia without denying the signal to the United States by transmissions at selective times.

In summary, the discussion of the availability of the NAVSTAR GPS signal and more particularly the precise accuracy is mostly a political issue as opposed to a technical one. If the Federal Government as a whole were willing to advocate NAVSTAR GPS as a national resource, a relatively simple solution could be found to allay the "scare story" that the signal could be denied to the civil sector. Alternatively, it is clear that detractors from the civil use of the system will be obtuse about such practical solutions. Satellites will eventually be used as the primary source of radionavigation. Denial is a minor issue, if it is a valid issue at all.

2.0 TECHNICAL CONSIDERATIONS

2.1 General

Technological advances during the past several years are providing the capability to implement a satellite-based navigation system concept with enormous potential benefits to aeronautical, maritime and terrestial civilian users. The NAVSTAR Global Positioning System represents a potential utility for an order of magnitude performance improvement for navigation and exploitable utility for precise time transfer, system synchronization, collision avoidance, and guidance.

The GPS satellite system concept is an outgrowth of numerous experimental and developmental satellite techniques and proposed implementations. References provide a detailed summary of the characteristics and features which have been pursued for satellite systems such as: TRANSIT, the Navy Navigation Satellite System; the TRANSIT Improvement Porgram, TIPS; the TIMATION and 621B development experiments; the NASA Position Location and Communication Equipment (PLACE) experiment; the Maritime Satellite Program of the Department of Commerce's Maritime Administration; the DOT/FAA Aeronautical Satellite Program; the DOT's Advanced Air Traffic Management System Concept (AATMS); and the FAA's ASTRO-DABS Concept.

2.2 Performance Characteristics

Consideration of the civilian application of GPS requires an appreciation for the basic system characteristics which have been incorporated into GPS due to military utility considerations. There have been many previous investigations directed toward establishing the navigation system characteristics which are considered to be important to the accomplishment of military missions. These military characteristics are listed in Table 1 and contrasted in a qualitative sense with civilian needs for navigation.

The NAVSTAR GPS concept to a large degree satisfies each of the specific minitary characteristics listed in Table 1. The desire to achieve these essential characteristics, combined with the limitations of existing system candidates, individual experiments, and the analytical development of navigation satellite capabilities dictated the GPS technological design. It should be noted that many items considered essential for satisfying the military navigation requirements have no similar civilian impetus or importance. This disparity exists mainly for characteristics of anti-jam immunity, selective denial, U.S. territorial ground control, and the need to be compatible and easily transitioned with respect to existing navigation aids. The degree to which civilian essential characteristics are compromised by the present GPS design would seem to be singularly in the area of potential obsolescence of existing navigation aids in which the civilian community has invested significant economic expenditures.

2.3 NAVSTAR GPS Capabilities

Satellite navigation as embodied in the GPS concept is characterized by line-of-sight operation at long distances between the users and the space satellite vehicles. To indicate the technical feasibility of providing the desired performance advantages of the concept, each of the significant characteristics need to be considered.

2.4 Worldwide Coverage

The operational GPS satellite segment will employ 24 satellites in 12 hour circular orbits. These satellites are equally spaced in three orbital planes so that six to eleven satellites are always in view to a user on or near the surface of the earth. Average visibility for all locations on the earth and for all times during the day is eight to nine satellites visible above 5^{o} elevation angle from the horizon, and six, seven or eight can be seen above 10^{o}. Since only four or less satellites are needed for civil navigation, the users generally will have twice as many satellites available from which to select the best geometry.

2.5 Spread Spectrum Code Modulation

The selection of a spread spectrum code modulation for GPS provides significant advantages for a satellite-based passive navigation system. Spread spectrum code modulation provides several essential system characteristics such as: Single frequency allocation for satellite transmissions

Performance and Adjoint Sensitivity Analysis of the NAVSTAR Global Positioning System", NAECON '76 Record, pp. 834-840.

29. K.A. Myers, R.R. Butler, "Simulation Results for an Integrated GPS/Inertial Aircraft Navigation System", NAECON '76 Record, pp. 841-848.

with code-division multiple access to each individual satellite, precision ranging measurement on the code modulation in continuous real time, significant processing gain against signal interference, and inherent multipath discrimination.

Other distinct features have been incorporated to benefit the potential user community. The GPS concept utilizes one-way ranging to the satellites with no active transmission from the users. This passive navigation approach can therefore support an unlimited number of military and civil users. A user employing an inexpensive time reference can be computationally synchronized to the GPS system time base by utilizing at least four sttellite signals from the global coverage, or only three if altitude or time are known. System time is maintained within a few nanoseconds at the GPS satellites by a Control Segment which tracks the satellite and determines the individual ephemeris and satellite clock parameters. All resultant navigation solutions are thus referenced to a common grid, the Department of Defense World Geodetic System, 1972.

2.6 Navigation Signal Structure

Navigation signals are transmitted by the satellites on two L-band frequencies; L_1 which is centered at 1575.42 MHZ and L_2 which is centered at 1227.6 MHZ. The L-band frequency provides all-weather operation with minimal ionospheric group delay propagation errors. Military users may employ two-frequency operation at L_1 and L_2 to precisely calibrate ionospheric delay. The signal waveform structure which is impressed upon the L-band carrier frequency is a composite of two pseudo-random-noise (PRN) phase shift-keyed (PSK) code signals transmitted in phase quadrature. These two signals are termed the precise or P-code, and the Coarse/Acquisition or C/A-code. The P-code provides a high precision navigation accuracy for military users which can be encrypted or altered to provide a secure code for selective denial purposes, and is also resistant to jamming and multipath distortion.

For the civil user, the C/A-code provides a ranging modulation which consists of a PRN sequence of binary "chips" biphase modulated on the carrier at a chip rate of 1.023 Mbps. As shown in Figure 1, each C/A PRN chip duration is equivalent to approximately 978 nanoseconds or a range duration of 293 meters. The C/A code has a 1023 bit linear gold code pattern generated by the module-2 sum of a selected pair of maximal PN codes from a 10-stage shift register generator. Each satellite has a unique C/A-code which is obtained by inserting a different time displacement (i.e., code address) between the maximal code pairs. The C/A code has a period of 1023 chips, or exactly 1 millisecond at chip rate of 1.023 Mbps. The repeating 1 millisecond long code duration provides unambiguous ranging equivalent to about 300 kilometers between the satellite and the user, which is compatible with the 10,900 nautical mile orbit altitude of the satellites. Gold codes of period 1023 are characterized by unwanted subsidiary correlation peaks which are 23.9 db below the main autocorrelation peak, consequently multiple access of all 24 satellite signals is feasible.

2.7 Navigation Data

Satellite system data is generated coherently with both the P and C/A-codes to define the individual satellites ephemeris and satellite clock data for basic navigation information. The satellite data frame also includes satellite identification, status, system time indexes, and telemetry words. The data is differentially encoded, non-return to zero, 30-bit words in 6 second frames generated at 50 bps. Total message data from each satellite consists of 30 seconds of data to define a 1500 bit data frame with 5 subframes of 300 bits. Each bit of the satellite data has a duration of exactly 20 msec. C/A-code repetitions. Time-of-day information in the satellite data frames conveys satellite time to the receiver with an accuracy of one data bit, or 20 msec. If the receiver locates the data bit transitions, the time accuracy derived from the data can be refined to less than 1 msec. which is the period of the C/A-code. This enables the receiver to establish unambiguous ranging to the specific PRN code correlation peak.

2.8 Navigation Signal Processing

The civil user receiver can selectively provide continuous ranging on a particular satellite by generating a local replica of the specific satellites corresponding Gold code.

The spread spectrum receiver performs several basic functions as illustrated in Figure 2. These are: Acquisition search to establish synchronization between the received PRN signal and the replica code; delay-lock code tracking to maintain synchronization and to extract the "pseudo range" measurement; carrier phase tracking to extract the doppler velocity or "pseudo range-rate" measurement; and data demodulation. Precise alignment of the replica code with the received satellite signal results in a well defined autocorrelation function.

The pseudo range measurement corresponds to the displacement of the autocorrelation peak with respect to the user clock referenced replica PRN code. When the received satellite signal is multiplied by a synchronized replica of the PRN code, the spread spectrum signal is despread and becomes a narrowband carrier biphase modulated by the satellite data. The ratio of the pre-correlation bandwidth to the post-correlation bandwidth is defined as the spread spectrum processing gain and quantitatively measures the interference rejection capabilities of the receiver. Conventional demodulation techniques for narrowband signals can be applied to the despread correlation output. In contrast to this, the pseudo-range rate is measured from the carrier frequency doppler, and this measurement is independent of the spread spectrum modulation.

2.9 Navigation Accuracy

The level of navigation performance envisioned for the Global Positioning System is a direct result of the selected waveform characteristics of the GPS satellite signal. Operation with either the P (10.23 MHz) or C/A (1.023 MHz) code modulations to a large extent establishes the basic navigation accuracy level which may be provided to the user. The system level navigation performance accuracy is therefore formulated by error contributor categories of P or C/A signals.

The other significant performance constraint on GPS is the environmental medium or propagation link in which the signals are transmitted and received. Performance in the L-band spectrum is influenced by natural phenomenon and by the mechanization techniques which are employed.

Error contributions have been allocated to the various system segment contributors; the Space Vehicle segment, the Propagation Link, and ultimately, the User Segment. These error contributions are shown categorically in Table II.

For each user observation of pseudorange toward a specific satellite the uncorrelated portion of the observed range error is termed the "User Equivalent Range Error" or UERE.

User navigation error defined in terms of three vector components of position and scalar system time is obtained by utilizing four independent scalar pseudorange observations. This resulting navigation error is thus defined by the UERE multiplied by the Geometric Dilution of Precision (GDOP) which is uniquely established by the geometric relationship between the user's position, and the specific positions of the four satellites utilized for the observations. These concepts provide an approximate means of defining navigation capability which is not associated with geometric orientations but which instead stipulates fundamental measurement or observational errors inherent to the navigation process.

The error budget for both P-code and C/A code UERE is given in Table III. The detailed error source descriptions and rationale for the budget is derived from information given in references. The prime difference between the military and civil user accuracies is the result of an increased ranging noise and the inability to compensate for the ionospheric delay.

Figure 3 summarizes the cumulative probability distribution function for the GDOP anticipated during the operational 24 satellite coverage phase. Navigation error in terms of both radial horizontal plane errors and vertical axis errors are indicated for a satellite elevation viewing restriction of 5^{o} above the horizon.

Using the data in Table III and Figure 3, the expected navigation accuracies for a C/A-code user are summarized in Table IV.

3.0 LAND

3.1 General

Throughout time, man has required information concerning the location of positions and objects. Such information is required by navigators, geologists, surveyors, mapmakers, aviators, air traffic controllers, etc. Accurate position location is mandatory for these and other uses. It is also mandatory to be able to locate land vehicles for purposes of dispatch and command and control.

3.2 Users of Land Position Information

Studies indicate that operators of land vehicle fleets can make their operations more efficient through effective command and control of individual vehicles. Such command and control decisions are highly dependent upon reliable and accurate information. Other users can not use the command and control capability as their fleets are small or they do not require this capability. These users require a simple, but accurate location system only. Depending on the type of operation, savings can accrue to a vehicle fleet operator through a reduction in the number of vehicles required to maintain a given level of service and through reduced operating costs by installing position location systems.

Among the potential civil land users of a location dependent information/command and control systems are the following:

Location information only -	emergency medical services
	fire departments
	delivery services
	repair servides
	utility services
	security and surveillance agencies
	geological survey services
	post office departments
	taxi services
	hazardous cargo transporters
Command and control systems -	police departments
	public transit.

As these uses fall into the general applications of location information only, and command and control systems, they can be discussed under these broad headings. Essentially, both categories require location information inputs. For location information only uses, manual use of the data is made

receiver's computer will require a capability to select the optimum four satellites (of a 24-satellite system) to navigate from at any given time. Finally, it must be able to process the data of the

* This work was sponsored by the Air Force Armament Laboratory, Eglin Air Force Base, Florida, under contract F08635-76-C-0343.

to record information or dispatch vehicles. For command and control application, the location information is fed into a computer which makes decisions based on the input location information. First, let's discuss the location only uses of position information. In the case of fire departments, location information can be expected to be only of occasional use. Generally, fire equipment remains at the fire station until it is assigned to an emergency (thus, its location is known prior to dispatch). Once a vehicle is dispatched to a fire and is enroute to the scene of that fire, accurate location information would permit a dispatcher to exercise a greater measure of control than is presently possible. Vehicles enroute to an emergency could be redirected to another emergency and vehicles returning from an emergency could be dispatched to another incident with a minimum of effort. Occasionally, fire vehicles converging on the scene of an emergency are involved in a collision between themselves. This has resulted in the loss of the vehicles and injury of the fire fighters. Accurate position location information would permit the dispatcher to alert drivers to a potential accident. The same is true of vehicle assignment of emergency medical services, such as ambulance and rescue services.

Taxi services may be improved through vehicle location information that is presented to the taxi dispatcher. Presently, when a request for taxi service is received the dispatcher assigns the call to what he believes to be the nearest vehicle who then responds. The closer the assigned taxi is to the request, the fewer the non-revenue miles on the vehicle will drive, thus resulting in lower costs to the operator. Accurate location information can assist the dispatcher assign the closest taxi to the request.

In recent years, a new form of urban transportation has been emerging in the United States. These are the random-route, demand-responsive systems that offer door-to-door service. A person desiring transportation telephones the dispatch center and requests transportation from his present location to his destination. The dispatcher assigns a vehicle that is near his location and is generally heading in the direction of his destination. Along the way, additional passengers are picked up and dropped off as the vehicle detours slightly from its course. Obviously, location information plays a role in vehicle selection and assignment. Small demand-responsive systems are able to be controlled by a single dispatcher who is able to approximate the location of each vehicle. Large systems use a computer to help in the vehicle assignment, however present systems rely on the dispatcher's estimate of each vehicle's location. Location information may be able to assist the dispatcher by relieving him of this chore.

Delivery and repair services likewise may benefit from vehicle location information. Examples of such repair services are public utilities and applicance repair vehicles. The operations performed by these services may be made more efficient if the location of the vehicles is known to a central dispatcher. A delivery service dispatcher, for example, can use location information to select a vehicle to pick up merchandise on the day of the request rather than schedule another vehicle to make the pick-up the following day. The same analogy holds true for utility and repair services.

Another use of precision on location information is for data recording. Data on population and site registration is presently collected by census bureaus, market research firms, highway departments, and geological exploration firms. Each of these users needs are largely self-explanatory. One major difference from the previously discussed requirements, is the need for this group of users to be able to return to the same spot using the position information.

Let's now discuss the application of command and control systems to the second group of potential users. Police departments typically field a large number of marked vehicles to patrol their jurisdiction in a random fashion as well as a smaller number of vehicles dedicated to investigation, towing, supervision, etc. Generally the jurisdiction is divided into a number of small areas referred to as beats. A police vehicle is assigned to one (or more, depending on beat size and conditions) beat/s and is expected to patrol in a random fashion. When an incident or emergency is reported, the police dispatcher determines in which beat the incident has occurred and radios that cruiser to respond. The present scenario assumes that the beat cruiser is closest to the incident. This may or may not be true. The cruiser may be patrolling an area of the beat which is far from the incident. A cruiser in another beat may be closer or another police vehicle may be passing through the area and thus be able to respond quicker. If the dispatcher knows the location of all police vehicles, the closest vehicle can be assigned to respond to an emergency. Studies indicate that dispatching techniques such as this may permit a reduction in the number of vehicles required to maintain the same level of service. And, of course, service can be improved using the same number of vehicles.

Bus transit operations can likewise benefit from centralized location information in the command and control of a large vehicle fleet. There are 11 U.S. transit operations with over 1000 buses. New York alone has 5000. Bus drivers are responsible for the on-schedule operation of their bus. However, due to conditions beyond their control, it is frequently not possible to operate within the allotted schedule. Traffic conditions, a fire or accident along the route, lighter or heavier than anticipated passenger loads, etc. contribute to off-schedule operation. Whenever one bus starts running late (or early), it impacts the operation of the entire route as well as affecting the service delivered to the public. Passengers are forced to wait longer for the bus and buses become over-crowded due to heavier loading of buses operating behind schedule. Location information presented to the central dispatcher will permit him to control the routes and thus significantly improve public service. Additionally, passenger and driver security can be improved if the location information is coupled to a covert, silent alarm that the driver can activate in the event of a criminal emergency. The alarm, once activated, will alert the dispatcher who will then notify the police, giving the bus location. Such an emergency location system, now in use at the Chicago Transit Authority, has reduced police response times to bus emergencies from a 7.15 minute average to a 4.26 minute average. While these times are based on limited

V_{LOS} = relative velocity along the line-of-sight between transmitter and receiver

C = velocity of propagation of radio waves

data, they indicate a *hopeful trend*. Additionally, the bus command and control situation can be automated by using a computer to exercise command and control over the bus fleet.

3.3 Present Use of Position Information

Land position information use generally falls into four categories: vehicle dispatch, schedule control, data recording, and emergency location. Each of these users has been discussed previously but one point remains to be made - location information must be relatively inexpensive to obtain, relatively accurate, and reliably available.

3.4 Benefits and Costs of Location Information

Essentially the benefits of location information result from improved operational efficiency, improved service, improved security (or perception thereof), and reduced costs of collecting data. These are not mutually exclusive but can be traded off between themselves. For example, a user of location information can improve the efficiency of his service and thereby reduce the number of vehicles used while holding the service level constant, or he can continue using the same number of vehicles and increase the service he provides. Thus, the benefits can be divided into three groups: economic, better service (that which does not translate into economic terms), and safety and security.

In the case of vehicle fleets, the economic benefits accure from increased fleet productivity and result from the number of vehicles which may be saved while maintaining a constant service level. These surplus vehicles are the catalysts which produce capital and operating savings. In the case of data recording, the economic benefits result from payroll savings. By permitting the rapid and accurate recording of data as well as assisting a person return to the location, lower payroll costs can accrue while surveys can be made more accurately.

Improved service benefits are difficult to quantify as they involve such intangibles as reduced passenger waiting times and more confident schedules. The difficulty in quantifying improved service benefits results from the controversy surrounding such things as the value of an individual's time.

Benefits from improved safety and security are likewise difficult to quantify. Such question as "What is the value of a human life?", "What is the value of preventing an injury?", etc. must be answered before the benefits side of the equasion can be filled in. Obviously, the monitary saving from preventing a truck hijacking can be readily computed by determining the value of the cargo that is delivered to its destination rather than directed to illegal uses. In any event, it can be readily understood that improved safety and security do have a value. Drivers fear and anxiety can be expected to diminish. Bus passenger anxiety can also be reduced and possibly new passengers can be enticed to ride the bus. Even though the statistical probability of a bus passenger being exposed to a valid threat is currently very small, improved security may alleviate unrational fears.

The Chicago Transit Authority has equipped its bus fleet with an emergency location system that is covertly activated by the driver when he is threatened. Other buses were equipped with a covert alarm that did not give location. In a limited test, based on 63 incident reports in which the driver specifically indicated police arrival times, alarm response times were reduced 40%. Following are the reported response times:

- Buses equipped with silent alarm and location	4.26 minute average
- Buses equipped with silent alarm but not location	7.15 minute average

From these average response times, it can be deduced that the lower response time may be critical in preventing harm from befalling bus operators and passengers. Almost three minutes response time was removed from the non-location equipped buses. These three minutes may be vital in a life and death situation.

3.5 Present Methods of Calculating Position

Present methods of calculating position on land are generally dependent upon manual input of location data. In some instances, as in the case of repair or delivery service, location information is frequently trasnmitted via the telephone. Transportation services, such as taxi and bus operators generally utilize a radio to transmit location information. Police location information is generally based on an assumed location (each patrol vehicle is assigned to a geographic district and is assumed to be in that district). Updated location information is manually (or vocally) transferred to the dispatcher who must then make a decision based on the information. The location information is accurate only so long as the vehicle does not move. Radio channel constraints limit the frequency with which the information may be updated and, of course, the human limitations of handling the data likewise minimize the amount of data the dispatcher can handle.

Several police departments and transit operators are experimenting and evaluating automatic vehicle monitoring systems that will automatically transmit location (and other) information to a central computer which will put it into format and present it to the dispatcher for action. Examples of such systems can be found in the St. Louis, Missouri, Huntington Beach, California, and Dallas, Texas Police Departments, and in the Zurich, Switzerland, Hamburg, Germany and London, England transit operations.

There are a number of location sybsystem types that can be used in automatic vehicle monitoring systems. Generally, they fall into four categories:

1. signpost,
2. radio frequency,

3. dead reckoning, and
4. hybrid combining two or more of the above.

The signpost location subsystem utilizes a reference device (such as a narrow-beam optical scanner, a microwave transmitter, magnets imbedded in the road surface or coded radio signals from a small transmitter, to provide location at a particular point. Within the signpost category, these subsystems can be either"sharp" signposts providing precise location only at particular points, or "broad" signposts which locate a vehicle within a broader zone. Among the claimed benefits of the signpost-type location subsystem are:

- each signpost provides only limited coverage, thus if one fails, system performance is not significantly degraded.
- each signpost is relatively inexpensive.
- as each signpost is assigned a unique identification code, the on-board receiver equipment can be relatively simple as the computations are minimized.

Among the disadvantages of signpost-based systems are:

- a large number of signposts are required to cover a large area. While each unit may be inexpensive, the cost for complete large area coverage may be high.
- as a large number of signposts are required, maintenance can be a chore simply due to the number of signpost installations.

Examples of the radio frequency location subsystems are the Global Positioning System, Loran, pulse tri-lateration, etc. As the readership is familiar with RF type location subsystems, no additional comment is necessary. Some of the claimed advantages of radio location subsystems are:

- they cover large geographic areas which other systems could not economically cover.
- generally, the transmitters have been installed by government users who maintain them, thus sparing the civil user a maintenance burden.
- the signal is generally available at no cost to the civil user, as the government operates it.

Some of the claimed disadvantages are purported to be:

- government operated systems are subject to being turned off in times of national emergency and thus civil use may be denied at times.
- the Signal-to-Noise Ratio in urban environments is highly variable, ranging from good to non-existant, and accuracy thus suffers.

A land-based dead reckoning subsystem uses a compass to drive direction of travel coupled to wheel odometers to derive distance of travel. Due to the natural "drift" of compasses and increasing error accumulation of odometers, dead reckoning AVM systems must periodically "reinitialize" to produce acceptable accuracy. One such system stores a map of the city in its computer and at each turn onto another street, automatically "reinitalizes" itself (the system assumes that a vehicle can only operate on streets and in alleys, which are stored in the computer).

Finally, a hybrid system can use a combination of the above systems to derive location. Obviously, sush a system would hope to combine the largest number of advantages with the fewest drawbacks of each system. For example, a dead reckoning system might use signposts strategically placed throughout the city to periodically "reinitialize" the system, or, a radio frequency based system might be combined with signposts to augment location inputs in those areas where the signal-to-noise ratio is so poor as to induce position error.

Once a civil land-based location system is installed, a number of uses readily come to mind. Virtually any large vehicle fleet operator may be expected to derive benefit from location information. For police applications, the vehicle closest to the scene of an emergency can be dispatched rather than assume that the "beat" car is closest. Closest vehicle dispatch can be assumed to minimize the response time of police vehicles and thus improve service to the public. Emergency medical service and fire vehicles (although they do not cruise on an assigned beat may likewise be candidates for such systems. It would seem that the chief benefit for these types of vehicles is at those times when they are returning to their station after an emergency assignment. As the percentage of time this condition exists is relatively small compared with their "available for assignment" time, their use of location information may be marginal and thus system cost may be a large determinate as to whether they are equipped with location systems.

Transit operators are very concerned over schedule control both for individual buses and for the transit system as a whole. The public expects buses to operate reliably according to a published schedule. The fact that bus operating times and schedules frequently do not coincide leads many potential passengers to seek alternate forms of travel. It is assumed that if bus schedules were reliable, additional passengers could be attracted to transit, thus resulting in greater utilization of the transit fleet and reduced dependence on the automobile. The U.S. Department of Transportation is developing an Automatic Vehicle-Monitoring system which will use location inputs (from a radio frequency signpost transmitter location subsystem) that will be automatically compared via a computer with bus schedules. Preprogrammed instructions and messages will then be automatically transmitted to the bus operator instructing him of actions to maintain schedules and headways. The system is being developed so that any location subsystem (which produces acceptable accuracy) can be used to generate bus location information.

Apart from vehicle dispatch and control, many additional civil uses of position information are envisioned. . Highway inventory surveys, site registration and data position fixing just suggest the possibilities. Highway departments have a great need for digital surveys of the highway network. Assume, for example, that a highway department wishes to reconstruct or repair a particular segment of road. The contractor is instructed to start and finish the construction work in terms of "X" distance from a known point (usually a highway marker). Accuracy thus depends upon the location of the marker and the accuracy of the measurement. Also, highway data filed in this manner is difficult to store in a computer. Accident sites are similarly recorded. A police accident report frequently lists the accident site relative to a highway mileage marker, road intersection, or town boundary. Again, data such as this is difficult to code for processing and storage in a computer. Census information, forest fire location, etc. face similar problems. Thus, some sort of uniform position location information location is required.

3.6 <u>Requirements for Automatic, Precise Position Data</u>

For any positioning system to be widely adopted by the civil community, several requirements must be met. The requirements placed on any system for automatic precise location data can generally be considered to be:

(1) Reliability - the system and its desired information with minimum maintenance costs.

(2) Availability - the signal must be available to all users and provide wide-area coverage (coverage is defined as the operating area of the user).

(3) Communications and Frequency Availability - the location data must be formated so it can be transmitted to a central computer efficiently (in the case of vehicle location).

(4) Accuracy - the location information must provide accurate information concerning positions.

(5) Cost - the location information must be affordable both in terms of capital and operating costs.

Reliability is as important to the civil user as it is to the military user. While reliability is frequently defined in varying terms, several important measures are "System availability expressed as a percent of Total Time" and "Mean Time between Failure". If the position information is being used by a police department, reliability is important as lives depend on continuous position updating. If the position information is being used for geological exploration, reliability is important as considerable expense is incurred in placing a survey team in remote field locations. Each user will define the reliability he needs and in many cases they will vary. Generally, however, system failures must be minimal, and the user must be confident that his reliability requirements are met.

Availability is equally important to the civil user. Accurate position information must be available to all civil users without burdensome review of his application. It would be unfair to permit some users unrestricted access to data from positioning systems while denying it to others. One potential conflict of this nature would arise if civil governmental use were permitted, while civil non-governmental (e.g., delivery services) were denied use.

Another aspect of positioning system availability is that it must be available to the civil user year-round on a 24 hour basis. Under no circumstances may the system be turned off or denied to the user. If a police agency, transit operator or geological survey team places the success of its operation in the hands of an automated command and control system or data collection system using position information, then chaos could result if the position information were suddenly denied them. Police vehicle deployments and transit operations would become disrupted. The geological survey team might not be able to accurately pinpoint the location of natural deposits previous efforts have uncovered or to be able to return from remote sites. The point of this is that if the position location data is government supplied and becomes widely adopted by the civil sector, then the data may not be denied in times of national emergency for it is precisely during such times that police, transit, geological exploration, and national resources recovery assume paramount imporntance.

A third aspect of availability is coverage. Whatever position location system is adopted by the civil user, coverage of his service area must be assured. The coverage area of a police department can range from a few square kilometers to millions of square kilometers (as in the case of a national police agency). Agencies responsible for nuclear shipments also have a requirement for national coverage so they can track shipments across the nation and assure the safe arrival of these shipments. Generally, the greater the coverage area a position locating system provides, assuming the system's accuracy meets the user's needs, the greater the number of potential users the system will appeal to.

Communications and frequency availability are important for any position locating system's success in the field of civil vehicle command and control. Command and control systems require real-time data to operate. For a command and control system to be effective, radio frequencies must be available to transmit data. Thus, communication frequencies must be available for implementation of command and control for vehicle surveillance systems to be widely adopted. These frequencies are becoming increasingly difficult to obtain, especially in heavily populated urban areas. This has lead to a trend of adding additional vehicles to radio channels with the subsequent problems of channel congestion. Position information is a likely candidate for such transmission as digitally formated position information which can be transmitted in less time than if transmitted by voice. To shorten the digital

transmission as much as possible, some amount of data processing must take place on-board the vehicle. Raw data from a position locating system can be shortened by using microcomputers in conjunction with the on-board equipment. Even with this processing, a substantial amount of information must be transmitted. Such transmission would probably include some message code, the vehicle identification number, position information, passenger counts (in the case of a transit vehicle), and any status messages the operator/driver wishes to include.

The accuracy required of any position identification system will vary with each user. Some users require high levels of accuracy in the order of 10-15 meters. Others require 95 to 100 meters accuracy, while others require only 150-200 meters. For example, geological exploration may require very precise location records so exploration team can explore the site a survey team has identified as promising. Searching for scarce, narrow-vein mineral deposits places high accuracy location requirements on the survey team. Police departments and transit operations require a lesser accuracy for command and control/dispatch operations. Studies have shown that accuracy on the order of 95 meters is sufficient to capture the benefits associated with command and control system operations. Significantly better accuracy produces only marginal improvements in system effectiveness while relatively poorer accuracy results in significantly reduced system effectiveness. Similar accuracy appears reasonable for the urban truck hijacking scenario. Other potential users, such as utility vehicles, could use somewhat lower accuracy, on the order of 200 meters, as their dispatching effectiveness is not as highly dependent upon accuracy.

The accuracy of radio frequency-based location systems is influenced by the signal-to-noise ratio (SNR) experienced by the receiver. While one can usually expect an acceptable SNR in rural areas, the urban environment severely degrades the SNR. Natural and man-made noise induces a severe handicap on all RF dependent location systems. Automotive ignition noise, buried and overhead power lines, etc. magnify the natural background noise conditions commonly found elsewhere. Added to the noise problem, multipath is very prevalent in the urban environment as is signal shading by high rise buildings.

Mutlipath is the condition in which signals are bounced from structures and multiple signal reception is recorded. A receiver must differentiate these multiple signals if accurate position information is to be determined. Often, "shading" of the receiver by the canyons created by buildings induce another source of error into the position calculation. Obviously, these sources of error must be overcome if a location system is to be adopted for the urban application.

3.7 Use of NAVSTAR GPS for Location Information

NAVSTAR GPS represents a bright star for future position location. When fully operational, it should permit world wide accurate position fixing. It may completely replace or supplement existing location systems in the land situation. In either event, it may provide a valuable service. The level of acceptance for GPS will depend upon several factors: its cost, both capital and maintenance; reliability, both in terms of equipment and signal availability; accuracy (relative to cost) and the availability of communications to transmit location information.

For GPS to supplement existing location systems it must match the capability and costs of existing systems. Assuming the match is about equal, new users may be attracted to GPS. For existing location information users, there may be no clear advantage to adopting the newer GPS system. Obviously, in those locations where the GPS signal is the only one available, there is no alternative but to adopt GPS.

Potential market size is always difficult to estimate as it is so cost dependent. The potential market (in number of units) for U. S. civil land use of location identification systems is estimated to be approximately:

police vehicles	170,000
fire vehicles	200,000
emergency medical vehicles	28,000
urban transit vehicles	50,000
interstate buses	20,000
intercity trucks	1,000,000
delivery vehicles	1,200,000
utility company vehicles	100,000
taxis	50,000
census inventory	50,000
land transportation statistics and inventory	10,000
geological survey and exploration	2,000
Total estimated market size	2,632,000

As is evident, the market for civil land use of location information systems is substantial. Whether a particular type of system can capture the market is heavily dependent upon a number of factors, many of which have been discussed. The system must be reliable, easily maintainable, available and affordable.

Assuming the answers to questions concerning the above characteristics are positive, cost will assume major importance. There is no single answer as to what each user is willing to pay for a loca-

tion information system. An organization responsible for the safe and secure transportation of nuclear shipments can be expected to be willing to pay more than lower priority users. Intercity turcks transporting high value cargo which can be readily disposed of (such as liquor or cigarettes) may be prime candidates for such systems, while carriers of low value bulk goods (such as sand and gravel) may have little use for electronic systems.

The market then, is relatively sensitive to cost. It is probably safe to speculate that full command and control systems costing more than $4,000 per vehicle (including all central control displays and consoles, software, and in-vehicle equipment such as receivers) will probably have a difficult time in the civil land use market place. For the position location only market the receiver and data recording mechanism should cost less than $1,500. This cost estimate assumes that the receivers will cost approximately $1,000 and is equal in price to LORAN-C receivers that are expected to appear on the market shortly.

3.8 Conclusions

There is a definite, yet relatively undefined, market for position information and command and control systems in the civil land-use market. Certain potential users, such as police, urban transportation operators, transporters of high-value and hazardous cargo, and geological survey organizations can be the initial users of such systems. Gradually, as experience with these systems develops, additional uses will be identified and the potential market size will inevitably be increased. As the market size increases, it can be assumed that production economies of scale will result in lower equipment costs, thus further expanding the market.

4.0 SEA

For milenia, men have used Watercraft to ply routes of trade throughout the world. In early times, man's knowledge of the world was as limited as his ability to produce, transport, and sell goods. His progress at sea was similar to his success in growing crops and capturing game necessary for livelihood dependent upon whims of the gods! As man's knowledge of his earth and universe increased, so did his sophistication as a seafarer.

Along with his sophistication, the Mariner developed a fierce independence from ordinary land lubbers and their clumsy assistance. With months, even years, at a time to become at-one with the sea and the sky, he became as familiar with the sun and the stars as landfolk were with their own living quarters. The Mariner recognized the hazards of his life and treasured the rewards of his independence. Until a little over a hundred years ago, wind and current, aided rarely by strong backs of men provided the only motive power for ships.

4.1 Technology Revolution

Steam engines and iron hulls thrust themselves into maritime commerce in the late nineteenth century. The old heritage and traditions of the sea were slow to give way. Innovation was looked upon as sacriledge and adopted (often unaccepted) only when law or regulation forced new ways upon the reluctant seaman.

4.1.1 Aids to Navigation

Commerce on the high seas was accomplished by the wiles of men and the vagaries of nature with little assistance from technology. Buoys, lighthouses and daymarks were the only aids to navigation available or used to any significant degree until the first quarter of this, the twentieth, century.

4.1.2 Radio Aids to Navigation

Until the invention and use of the radio the independence of the Mariner, with his sextant, was a necessity. Prominent disaster in the sinking of the Titanic and agony and catharsis of the first world war were necessary to attract the attention of mariners toward the capability of radio to improve his lot. Radio Direction Finders became the only major tool for the navigator.

An explosive growth of radio, electronic and associated technologies was forced by the cruel necessity and frightening exigencies of a second Global War. In a short span of ten years the mariner had literally hundreds of radio systems for communications and navigation available to him. He did not see these systems as "manna from heaven" -- rather they were insults to his hard earned expertise, and intrusions upon the privacy of his lonely but proud way of doing things and of living.

4.2 Maritime Commercial Development

In the less romantic, but nevertheless demanding, economic areas of maritime commerce, ship owners and operators have made staggering progress. In terms of tonnage a present day single 480,000 ton tanker carries the total tons of overseas maritime commerce conducted in one year by the United States in 1801. In 1900 U.S. trade was a million tons. In 1975, total trade was 403 million tons. By the year 2000 World Maritime trade is expected to be between two and four Trillion tons.

4.2.1 Radio Navigation Systems Status

Table V shows major existing radio navigation systems with attributes to commend them as well as attendant deficiencies. Most of these systmes were spawned by wartime or military development and implemented very slowly in the maritime community. Only Radio Direction Finders and recently RADAR have crept into the regulatory language of the international maritime community.

Radar and the fathometer have given the mariner "eyes in the dark" and removed much of the hazard of fog. The fathometer and sonar allow close range underwater perspectives.

Loran and Omega are in the vocabulary of a fairly large number of mariners. Loran provides precision in areas of critical need. Omega provides continuous general service where moderate services and accuracy are sufficient.

Satellite capability to tie the world together on a common geoid and to provide at-sea periodic precise location on that geoid has been demonstarted in the Transit system. Several other experimental systems have further demonstrated satellite potential.

4.2.2 Requirements

In spite of all of the amazing sophistication in many areas, the maritime community still relies on the simplest of devices to assist a pilot. Handling the awesome size and mass ot today's superships still depends largely on the acquired skill of the pilot. His senses of sight, sound, smell and motion remain the most reliable and widely used tools of his trade. These senses are the ultimate arbiter in his decisions. Staggering catastrophe for port areas and entire shore environments is the consequence and judgement for his errors.

With the realization of these potential consequences, it must be recognized that Maritime Requirements are not well recognized.

4.2.2.1 Mariners Perspective

The mariner depends first on his senses, next on the lore of the sea, then on conventional visual aids, and finally on electronic devices. Redundancy in systems is of primary importance. His acceptance of aids is generally proportional performance pressures. On the high seas, with many fathoms of water under the keel, he readily accepts and uses any and all radio aids available. In heavily trafficed areas and costal confluence regions, he readily uses radar to assist him when his visibility is restricted. In estuaries and harbors he is least willing to use external systems, and is most demanding of redundant capability.

4.2.2.2 Who is the Mariner

In order to gain some perspective of the Mariner, it is interesting to find out who he is, what he does, and how many of whom do what. In Table VI the different kinds of marine activities are shown along with estimated distribution. The Mariner's navigation requirements obviously will vary according to the purpose and type of each class of activity.

4.3 The Mariners Requirements

The mariner is terrified of the spectre that, if he states a desire for any service, a law or regulation will immediately be imposed requiring that he make provision to acquire equipment for and utilize that service in plying his trade. A few forward looking associations representing segments of the Maritime Community have made bold statements -- mostly about what might be required by other groups of seafarers.

4.3.1 Surface Fishing

Tuna fishing is discussed to illustrate this type of activity. Once a school of tuna has been located (often, unfortunately, by following the porpoise) boats may follow and fish based on visual observations of smaller fish activity in panic to escape the tuna feeding at or near the surface. Between sunset and sunrise the tuna sound to great depths. This fortunate characteristic gives a respite to the fisherman. Another characteristic of tuna sounding is their almost infallible morning surfacing in the location where they sounded the previous evening. Accordingly, the tuna fishermen's navigation requirements are based almost entirely on the habits of the species for which he fishes.

4.3.2 Bottom Fishing

Several types of bottom fishing define different requirements for navigation. Shrimp are harvested by dragging nets along a smooth bottom. Discovery of an uncharted wreck or outcropping could ruin a day's fishing or perhaps an entire fishing trip.

Other types of fish live in isolated bottom rocks, wrecks, or other convenient hiding places. Location of these hideaways on a reliable basis week-to-week or year-to-year is essential to fishing success.

4.3.3. Shoal Fishing

Some classes of fish feed on the lee side of rocks shoals or obstructions. For success in fishing for this type, it is necessary to determine the location of the shoal and current direction at the edge of the shoal or obstruction.

4.4 The Outsider's Perspective

For those who are engaged in other transportation industries on land or in the air, it is difficult to comprehend the Maritime Area. Particularly in aviation, there is a discipline of communications, navigation, and interactive advice and control which is readily accepted and utilized continuously.

Indeed, without such organization and discipline, modern aviationcould not exist. How, then, could such an advanced system of maritime transportation been developed and expected to expand without imposition of restraints for commercial control and/or government intervention and regulation? Certainly the occurences of massive catastrophic oil spills with attendant contamination of fishing grounds, shorelines, and destruction of affected ecological life systems crystalize the conclusion that some surveillance and control is essential.

4.5 NAVSTAR GPS System

The GPS seems to have arrived on scene at a critical time in the development of maritime commerce. *Its appearance is nearly coincident with the recognition of global requirements for new concepts* of surveillance and control in maritime activity.

The GPS has the capability of providing services eclipsing all of those from other predecessor navigation systems, with essentially none of the disadvantages of limitations of earlier terrestrial (or satellite) systems. Attendant with the arrival of GPS is the near term availability of necessary communications support systems. The U. S. Marisat system is but a precursor for a new international INMARSAT system.

4.5.1 Potential Applications

In considering the potential applications for GPS, it is important to understand the motivational influences for a user to acquire new capability. Some of these factors for the Maritime Community are listed in their order of importance:

(1) Regulatory requirement.
(2) Direct replacement of older equipment essential, and used daily for operation or function-based on
 - a. Personal experience
 - b. Experience reports from others
 - c. Peer pressure.

(3) Favorable comparative benefit/cost assessment.

Although a regulatory requirement ranks first in importance to bring new capability for a vessel, often the performance available from equipment is the minimum acceptable, provided at the lowest cost.

The second influence, replacement of a proven performance, is a much more constructive motivation, but is much more difficult to approach.

The third influence is that most often used by new equipment/system salesmen, is that least likely to be successful in the operating maritime community. The benefit/cost of an equipment or system becomes apparent first in the financial management areas. In a medium to large company where management is separated from operations, there is usually an established distrust, if not active dislike, between "operators" and "comptrollers".

With these constraints in mind it is now appropriate to approach potential applications of GPS in the community. With its splendid capabilities it is evident that applications for GPS transcend application of all other navigation systems. With this recognition, it is now simply a marketing problem which will be assessed in the following paragraphs.

4.5.2 Marketing Strategy

With the currently increasing number and severity of environmental catastrophies which mar out environment and excite widespread public attention, the capability of GPS as a precise global navigation system will commend GPS utilization to meet regulatory requirements. In this context there are other conservative forces which will mobilize to oppose the adoption and spread of GPS almost as if GPS were a deadly disease. These conservative forces are governments and manufacturers of older, less capable systems and/or user equipment for these systems. These forces see, from a narrow viewpoint that GPS is a threat.

GPS marketing strategy must include an approach to these conservative influences which is convincing of near term special interest benefit as well as longer term general public benefits. Indeed if such specific benefits are sufficiently apparent, elements normally in general opposition can be converted to firm allies for mobilization to sell GPS.

4.5.3 Maritime Market

An essential part of Market Development is identification of the current market and projection of what the future market will be. Although the U. S. market is not entirely representative of the global market for GPS, studies made for assessment of the future U. S. requirements are available and are considered indicative of the future in the maritime industry (see references).

4.5.4 Market Penetration

As a general rule, the GPS penetration to be expected in a particular market area can be expected to be on the basis of the marginal profitability enjoyed by a particular user sector. With this in mind, cargo demand becomes controlling over other factors in determining marketability of a support service such as GPS. Therefore, it can be anticipated that high-demand cargo, transport-crude petroleum, or petroleum products is the user sector, initially most receptive of GPS sales efforts.

This sector is incidentally the group most likely to be targeted in regulatory action. Penetration in other market sectors will be dependent primarily on conventional market pressures.

4.5.5 Market Elasticity

Elasticity of a market is generalized as quantity of additional sales as a function of unit price of units to be sold (user equipment in the GPS case). An inelastic market is defined as a market in which unit sales are not affected by unit price. Negative elasticity is characterized by increased sales as unit price increases. Positive elasticity is characterized by increasing sales with decreasing unit price.

Production cost elasticity is a factor which, in the business world, must be evaluated along with price elasticity. Although production costs are not the only factor in delivered unit price, these costs are generally trigger values which influence an investor to mount a marketing effort to sell "x" number of units at a price of "y" dollars (marks, yen, pounds, etc.) Production costs generally are closely related to quantities to be fabricated in a production "run". For consumer goods such as television sets, the factor of time to sell "x" units at a price of "y" often determines whether a venture is profitable.

Although GPS probably will never become equatable to consumer devices, production runs of a few thousand can be reasonably undertaken to meet demand in particular market sectors.

4.6 User Population and Accuracy Requirements

Table VI sets forth a list of marine population and radionavigation requirments or accuracy and coverage. The table shows accuracy requirements by documented user class, including accuracy requirements versus area of operation. Areas considered are High Seas; Coastal Confluences Zone; and Harbor Entrance Zone. Operations are indicated in the areas by (-). Operations overlap areas in many cases and user populations are available only in terms of totals, therefore accurate distribution of the same user class between areas may not be possible and is only indicated as an overlap.. Total world vessel prpulation over 100 G.T. is shown; number of U.S. vessels documented over 5 G.T. are also tabulated.

Generally the requirement other than accuracy are the same throughout all areas as follows:

(1) Coverage: Coverage is required where there is commercial traffic, e.g., established ocean routes, throughout the CCZ, and in major ports.
(2) Availability: Continuous availability is desire.
(3) Capacity: System capacity to serve the user must not be exceeded.
(4) Ambiguous position: There must be no ambiguity.
(5) Cost: All require low cost user equipment.

The accuracy requirement for the high seas is 4 NM, 95%; for the CCZ it is 1/4 NM, 95%. A universal accuracy requirement cannot be stated for the HEZ. Visual aids provide for safety of navigation. If an electronic aid were to be provided for use in place of visual aids, the accuracy as a minimum should be as good as that provided by the visual aids. Special users in all areas may have more or less stringent requirements.

4.7 Conclusions

The "bottom line" in assesment of system utility is user acceptability. In the maritime community to date, all equipments manufactured for use with all radio navigation systems (except radar, fathometers and direction finders) number in the order to two or three hundred thousand. The largest production runs undertaken number in the range from three to five thousand.

In the United States, the Maritime Administration has assumed a role in the GPS project. The objectives of the Maritime Administration are to:

(1) Assess applicability of GPS to meet current and future U.S. Maritime requirements.
(2) Focus various Maritime Sector need to "agglomerage" a user market.
(3) Promote maximum commonality in user requirements to obtain minimum user costs for equipment.
(4) Coordinate and define extended requirements for GPS to support National (or Global) objectives for environmental protection, safety and other vital interests.

5.0 AIR

5.1 General

From an operational civil aviation point of view, the GPS should be considered as an Area Navigation (RNAV) system. The Federal Aviation Administration's Advisory Circular 90-45A (in process of being updated) sets forth basic considerations involved in introducing RNAV into the National Aviation System (NAS). At present, the great majority of the RNAV systems in use are based on VOR/DME (VORTAC). (Figures 4 and 5).

Other navigation systems which currently may be considered to have RNAV capability include Loran C, VLF/Omega, TACAN and INS/Doppler. These systems may be approved by the FAA for Instrument Flight Rules (IFR) operation enroute, in terminal areas and for instrument approaches provided they equal or exceed the VOR/DME RNAV accuracies as specified in AC 90-45A. (Figure 6) So far, none of these other RNAV systems have been approved by the FAA as a primary IFR RNAV

system.

Because of the angular divergence of the VOR radials, RNAV route widths may be in excess of 4.0 nm on either side of the route centerline in an angular splay, depending upon distance from the reference facility. Inasmuch as VOR/DME facilities are subject to line-of-sight (radio horizon) limitation, RNAV instrument approaches are not authorized at locations more than 25 nm from the reference facility. Under the best of conditions, RNAV instrument approach minimums generally are not less than MDA (Minimum Descent Altitude) 400' and visibility 1 nm (Figure 7). Helicopter visibility minimums may be half of those approved for fixed wing aircraft instrument approaches. All RNAV instrument approaches are classified as "non precision". A "precision" approach according to current FAA definition can only be made with an ILS (or future MLS) or a precision approach radar (PAR/GCA) at the point of intended landing. A Category I precision instrment approach facility provides minimums of 200' and $\frac{1}{2}$ nm; Category II 100' and 1200' RVR, and Category III in three gradations down to 0-0.

5.2 RNAV Categories

There are three basic categories of RNAV systems:

(1) 2-D - Whereby the pilot may determine his cross-track and along-track position (x-y coordinates).
(2) 3-D - Whereby the pilot may determine his x-y coordinates plus his z (altitude) coordinate in relation to a desired vertical profile (also referred to as "VNAV"). (Figure 6).
(3) 4-D - Whereby the pilot has the ability to arrive precisely at a point in space at a desired altitude on a desired track and at a desired time (x-y-z-t) coordinates). 4-D RNAV also permits arrival at a desired touch down zone (TDZ) on an airport or heliport at a desired time. (Figure 9).

At the present time, most general aviation RNAV systems are 2-D. Airline systems generally are 3-D. 4-D systems are still in development, but it is doubtful that they can be made sufficiently accurate using VOR/DME sensors.

5.3 Ideal Navigation System

In the present NAS, precision approach capability is provided at only about 500 civil airports and yet there are over 13,000 airports and heliports in the United States and its possessions. In addition to general aviation's needs for "precision" instrument approach capability at thousands of conventional airports, there is even greater instrument approach "precision" capability required when considering the growing need for precision approaches to virtually an infinite number of landing/takeoff areas for IFR-capable helicopters and other VTOL's (vertical takeoff and landing vehicles) now being developed. For example, approaches to city center heliports, to discrete helipads at conventional airports, to constantly changing locations for servicing of pipelines, to thousands of offshore oil platforms, to unpredictable locations for emergency rescue.

The ideal navigation system would be one which would have all of the following capabilities in an integrated system:

(1) Highly accurate enroute navigation accuracy so that airway/route widths could be in the order of $\pm$ 0.5 nm (linear) either side of the centerline.
(2) Sufficiently accurate approach and landing guidance so that "precision" instrument approaches could be made to any pilot-selected point on the surface without the need to have an electronic landing aid at that location.
(3) Ability to function without line-of-sight (radio horizon) limitations down to and on the surface.
(4) 3-D navigational guidance without the need to rely on any form of barometric altimetry.
(5) Highly precise time referenced navigational capability.
(6) Imperviousness to atmospheric conditions for noninterrupted operation.
(7) Non-saturable capacity.
(8) Service availability to all classes of airspace users on a worldwide basis.
(9) Cost effectiveness based on life cycle cost analyses, with system design such that it can have various levels of sophistication and thus will be affordable to all classes of airspace users.

It is considered that the GPS could well meet all of the above goals.

5.4 RNAV/GPS Implementation

On the basis that GPS will be used as an RNAV system for civil aviation, it is interesting to compare FAA RNAV and DOD GPS respective implementation planning. This comparison is shown in Table VII based on currently available data. (Note: Since issuance, the FAA RNAV implementation time table has slipped considerably, as have the GPS implementation phases.)

It will be noted that GPS RNAV could fit into FAA RNAV implementation quite smoothly. A study is needed to deal with the general subject of GPS RNAV interaction with FAA NAS planning in more depth, especially taking into consideration the implications of the latest FAA RNAV implementation policy announced in January of 1977, which in effect states that RNAV will be the navigation system of

future in the NAS.

RNAV as an operational function of GPS bears significantly on both GPS equipment design and future acceptance into the NAS. The foregoing indicates that the introduction of GPS RNAV could be accomplished smoothly and on an evolutionary transitional basis.

5.5 GPS Interaction with FAA E&D Programs

Another factor which bears on both GPS design and NAS acceptance will be the interaction of GPS with the FAA's Engineering and Development (E&D) programs, which are based essentially on the so-called Up-Graded Third Generation (UG3RD) ATC System concept.

The FAA UG3RD E&D planning can be broken down into 21 distinct and identifiable programs. The Table which follows sets forth each of these programs, with an "X" to show items that will be impacted by GPS to one degree or another.

A study is needed to describe the GPS implications for those items of the FAA's E&D programs which will have a significant impact or bearing on the GPS receiver and system design concept/s.

5.5.1 E&D Program Plans

Item	Subjects	GPS Impact
1.	Design Concepts for the UG3RD ATC System	X
	Airport Capacity Study	X
	Intermittent Positive control	X
	Automated Terminal Service	
2.	ATC Surveillance Radar	X
	Data link surveillance	X
3.	Technical Development Plan for DABS	X
	ATC Radar Beacon System Plan	
4.	Navigation	X
	Area Navigation	X
5.	Airborne Separation Assurance	X
6.	Communications	X
	Data Link Operational Experiments	X
7.	ILS Improvement Plan	X
	Development Plan for MLS	X
	All Weather Landing	X
8.	Airport Surface Traffic Control	X
	Airport Pavement	
9.	Airport/Landside (not currently included)	
10.	Oceanic Air Traffic Control Automation	X
11.	ATC Systems Command Center Automation	
12.	Enroute Automation	
	Enroute Control	X
13.	Flight Service Station	
14.	Terminal Tower Control	
15.	Weather	
16.	ATC Technology	X
17.	Satellite Experimentation	X
	Aeronautical Satellite Communications	X
18.	Aircraft Safety	X
19.	Aviation Medicine	
20.	Aircraft Propulsion Systems/Air Pollution	X
	Aircraft Noise and Sonic Boom	X
21.	Aircraft Wake Vortex Avoidance System	
	Performance Assurance	X

A number of functions outlined in the foregoing table have data-link aspects which, when coupled with an airborne GPS receiver, can drastically alter present ATC System planning and provide many NAS benefits. For example, aircraft position in precise terms (x-y-z) could be automatically sent via data link to the cognizant ATC facility, thus providing a surveillance capability supplementing or in lieu of radar surveillance. Automatic exchange and comparison of x-y-z coordinates between aircraft via data link could provide a new dimension to collision avoidance (separation assurance). These and other data link applications to GPS for civil aviation should be examined further with due regard to receiver and ATC System design consequences. (Figure 10).

5.6 *Displays*

Since the primary operational function of a GPS receiver will be to serve as an RNAV system, a suitable control and display unit (CDU) will be needed to interface between pilot and receiver.

As mentioned in section 5.4, it will be relatively simple to integrate the GPS system with the FAA's RNAV planning. Since waypoints are used to define a desired RNAV route, including an RNAV instrument approach, the waypoints may be expressed either in rho/theta with reference to a VOR/DME facility, or in terms of lat/lon. The FAA is now well underway in identifying RNAV waypoints in both

rho/theta and lat/lon.

Since the GPS receiver will use lat/lon inputs to define waypoints, a GPS CDU will need to be designed to receive lat/lon inputs (similar to INS). The CDU can be designed in various degrees of sophistication, such as different capacity levels of waypoint storage; 2-D, 3-D or 4-D operational application; parallel track offset; time to go; ground speed, and so on. A study should develop various GPS CDU concepts with relative design pricing. (Figures 11, 12 and 13).

The simplest cockpit navigation display unit will be today's CDI/HSI. For track guidance, the vertical track bar will be used and for vertical guidance (if a 3-D GPS system is used) the glideslope horizontal bar or bug (Figures 14 and 15). More sohpisticated CRT displays can be included, along with preprogrammed flight data storage units. To the pilot, everything will be the same in the cockpit of the airplane to which he is accustomed, except for the extraoridnary GPS navigational accuracy and the elimination of line-of-sight (radio horizon) limitations.

If time referenced navigation (4-D) is used, "T" display capability would be needed to reference desired time of arrival to a particular waypoint (in space or on the surface), together with a fast-slow indicator. It is contemplated that the GPS outputs could interface with such sophisticated systems as flight directors, autopilot, and auto throttle coupling. Its outputs also could provide improved cockpit instrumentation such as absolute altitude, and highly accurate rate of climb/descent and velocity.

As with the CDU input device, various levels of cockpit displays (output) need to be examined, with alternate pricing and operational scenarios analyzed. This phase of further study should include a dissertation on sophisticated CRT multi-function displays (MFD) on which many functions can be shown, such as the GPS RNAV tracks and waypoints, own aircraft position, other aircraft positions as a type of CAS (Collision Avoidance System), weather contours, and so on.

5.7 GPS Communication

An essential function for implementing any Air Traffic Control System is the requirement for providing communications. The GPS communications capabilities are considered to be:

(1) Provide information exchange using coded, non-voice digital signals or emergency voice.
(2) Provide data link for air-to-ground information transfer in the local ground, approach and departure, terminal, and enroute regions of flight sufficient to accomplish ATC surveillance functions.
(3) Provide data link for ground-to-air information transfer for the same regions of flight sufficient to accomplish ATC executive control functions.
(4) Provide data kink capability to perform air-to-air information transfer compatible with identification and collision avoidance requirements.

No existing or proposed communication data link system meets these requirements. The current communication function is accomplished by VHF voice channels operating with associated control centers such as the Air Route Traffic Control Centers (ARTCC's) and the Terminal Radar Control Centers (TRACON's).

The design of the required data link function also has two basic options to choose from in terms of providing coverage. These options are shown in Figure 16 and involve the use of a satellite configuration in the link to provide for over water transmissions not within line-of-sight and for communication to surveillance centers shadowed by terrain for low altitude users.

5.7.1 Time Division Multiple Access

Concerns for the communication function design are motivated by the necessity of conserving spectrum allocations and usage since only a finite frequency band is available to accomodate the numerous potential users and the large volume of exchanged data. GPS navigation has the singular utility of providing for the incorporation of a data link communication technique which is based on time division multiplexing, or Time Division Multiple Access (TDMA). The timing accuracy of the GPS provides a concept of TDMA with guard band times of one micro-second. This one microsecond timing compares with a required 600 microseconds per 100 nautical miles of propogation delay encountered in a one-way data link. Such highly accurate timing from GPS makes the TDMA appear to be most fruitful for pursuing. A planned military TDMA secure communication system, the Joint Tactical Information Distribution System (JTIDS) uses a similar precisely clocked time slot concept but is not available to civilian users. Generally, the time division concepts imply a very rigorously structured information format with one central master network control function. This constraint may limit the flexibility needed for handling future growth unknowns and for accomodating the numerous transient user terminals. Figure 17 depicts the GPS synchronized time division data link concept.

5.8 User Community

There are essentially four (4) basic user groups in the aviation community:

(1) Military (user potential covered elsewhere in this Agardograph).
(2) Air Carrier.
(3) General Aviation Business (turbine and rotorcraft).
(4) General Aviation (private).

5.8.1 Air Carrier

The present number of U.S. air carrier aircraft is approximately 2,500. The figure is expected to grow to approximately 3,000 by 1982. An estimated growth rate of 4 percent is assumed after that time.

5.8.2 Business General Aviation

At present, approximately 10,000 aircraft exist in this category. This segment could grow to about 20,000 in 1982, of which about 5,000 will be IFR certificated helicopters. This group should increase by 5% to 10% after 1982. (Note: U.S. Coast Guard aircraft would be in addition to these numbers.

5.8.3 General Aviation (Private)

Present data indicae that approximately 130,000 aircraft fall in this group. The number should increase to about 190,000 in 1982 and then grow at about a 5% rate thereafter.

5.8.4 International Agreement

As found in the case of Omega, even the slightest idea of supporting a system which is basically military oriented, is generally opposed internationally. Having ICAO select a primary NAVAID generally takes quite a time period for negotiations (about 5 years) if past efforts are any indication. Stating that no GPS user taxes would be required, since its prime mission would be by the U.S. DOD (and NATO ?), would minimize this handicap. But, if user taxes are suggested, the air carriers and other civil users probably would not support the GPS approach internationally as the primary NAVAID.

5.9 Cost/Benefits Analysis

Four broad assumptions are made in the light of the foregoing dissertation:

(1) That the desired navigation and positioning requirements can be met most effectively by a satellite based system.
(2) That the satellite based system most likely to achieve these requirements in realistic time frame is the DOD NAVSTAR Global Positioning System.
(3) That the DOD under any circumstances would not deny use of the GPS C/A signals for civil aviation.
(4) That charges would not be levied against civil aviation for use of the GPS.

Thus, in order to determine the degree to which such a system would be used from the standpoint of civil aviation, a cost/benefits analysis would appear to be needed as the first step. As prior condition for such an analysis, however, a life cycle costing (LCC) analysis of the GPS would be essential. The civil aviation LCC analysis would use as a base line the DODLCC for military use.

Once this phase has been completed, an analytical sutdy should be initiated covering relevant aspects of civil aviation use of GPS including:

1. Development of GPS implementation and integration program plans.
2. Technical and operational system analysis.
3. Economic Benefit and Payoff analysis.
4. Specification of minimum acceptable GPS airborne equipment characteristics.
5. Generation of low cost GPS hardware.
6. Development of a detailed ground system phase-in plan acceptable to government and industry.
7. Test plan development.
8. Performance of test plan.
9. Analysis of test results.
10. Test evaluation and conclusions.

5.10 GPS Certification Considerations

5.10.1 Applicable Regulations

The Federal Aviation Regulations (FAR's) which bear directly on avionics standards for civil aircraft include:

- FAR 37 - Covers Technical Standard Orders (TSO's) procedures, particularly Subpart A which applies to administration of this FAR, and Subpart B which contains TSO specifications.
- FAR 23 - Covers certification of light aircraft (below 12,500 lbs. GTOW).
- FAR 25 (with Appendix) - Covers transport category aircraft.
- FAR 27 - Covers non-transport category helicopters.
- FAR 29 - Covers transport category helicopters
- FAR 127 - Covers helicopter scheduled air carriers.
- FAR 43 - Covers maintenance rules.
- AC's 43.13-1 and 43.13-2 - Cover acceptable means of compliance with FAR's including alterations.

5.10.2 Environmental Specifications

Radio Technical Commission for Aeronautics (RTCA) Document DO-160 provides the basis for FAA's environmental specifications. RTCA is a non-profit advisory organization serving the government (primarily the FAA), and consists of representatives of private industry and government, both

civil and military.

5.10.3 Certification

Avionics may be certified by the FAA in two basic forms: Technical Standard Order (TSO) or Supplementa Type Certificates (STC). (Note: a non-TSO or non-STC avionics unit may be installed in an aircraft by means of execution of a Form 337 which warrants that the equipment will perform its intended function within the intended environment.)

The issuance of a TSO is preceeded by an FAA Notice of Proposed Rule Making (NPRM) by which all interested parties may comment before the final rule (TSO) becomes effective. An example of TSO'd certificated avionics is the ATC Radar Beacon Transponder (TSO C-74c). The terms of an STC generally are set forth in an Advisory Circular, as in the case of AC 90-45A which covers STC'd RNAV equipment.

The RTCA prepares many documents on which the FAA bases TSO and STC standards. Others are originated in-house by the FAA.

A few questions which are pertinent in considering "minimum performance" for the purpose of establishing GPS certification criteria include:

- Is minimum performance required in visual weather (VFR) the same or different from minimum performance required when IFR in instrument meteorological conditions?
- Is the minimum performance required in aircraft carrying fare-paying passengers the same or different from that required in other civil aircraft?
- Is it practical to regulate and enforce the variables of installation and maintenance?
- Is it practical for the system to reward performance above minimum? What parameters suggest limits on maximum performance?

Optional equipment may be installed where carriage and use is a user option rather than a regulatory requirement. It must:

- Be designed, manufactured, and installed so as not to impair the airworthiness of the aircraft, and to perform its intended function.
- Comply with applicable requirements of FCC.
- Not deny or impair the service of any system element to other users such as through interference.

Required equipment must meet the foregoing three requirements imposed on Optional equipment, and in addition it must work to at least a specific installed performance minimum and possess minimum operational characteristics in order for other related elements of the system to function.

A study should consider the above points as well as the following elements pertinent to the establishment of design and related minimum performance standards for GPS airborne equipment.

5.10.4 General Considerations

- Establish the intended limits of use of the equipment or systems covered. The most comprehensive objective is to cover all applications. If, for some reason, objectives short of this are necessary to establish some critical element or application, the limited coverage should be explicitly defined. The needs of the users of all or portions of a report in this proposed format must be considered. Regulators must be able to adapt the document to their process, and producers and installers must be able to select applicable sections.

- Identify the test procedures to be described and explain their application and limitations. Such procedures should offer one or more methods of demonstrating compliance with performance standards, and must recognize that other methods not described may be acceptable. The categories of test considered in this format include:

(1) Bench Test - Tests in standard environment of equipment only. These should provide approximate guidance to equipment manufacturers for establishment of design objectives and monitoring manufacturing compliance.
(2) Environmental Tests - Tests in stressed environment for which the equipment was designed should be specified to assure performance in expected environmental extremes. Appropriate guidance should be provided for the installer or applications engineer concerning the environment in which operation is intended to insure that the environmental extremes for which the equipment is designed are not exceeded.
(3) Installed Tests - May be used in lieu of bench test simulation of such factors as electric power supply characteristics, magnetic field distortion, and interference from or to other equipment installed on the aircraft.
(4) Operational Tests - May include actual or simulated signal-in-space effects of antenna pattern anomalies, cable loss, power and error budgets, etc.

Any or all of the test procedures described may be used to demonstrate compliance with a particular standard, but only those tests essential to the purpose should be required. Reports should use only the simplest test procedures essential to establish compliance with minimum operational performance standards.

5.11 Conclusions

5.11.1 It is considered that GPS is an excellent candidate for a primary civil aviation NAVAID in the future. It has a number of advantages, some of the salient ones being as follows:

(1) Worldwide coverage.
(2) Excellent accuracy both 2-D and 3-D.
(3) Relatively immune to atmospheric disturbances.
(4) Unlimited capacity and signal coverage down to and on the surface.
(5) Relatively small O&M costs (few ground stations).
(6) Satellite redundancy results in a fail-safe capability.

5.11.2 But a number of factors must be resolved:

(1) Range/power/geometry/antenna problems must be considered since avionics costs are affected.
(2) Multi-path problems (shadowing) must be resolved if terminal area applications are considered.
(3) Utility and cost-effectiveness for the civil user must be demonstrated.
(4) International approval (ICAO) may be difficult due to its primary military role.

5.11.3 Assuming that these disadvantages could be overcome, it is believed that the GPS technique is an advantageous replacement for the VORTAC system by the end of the Century. Its worldwide coverage, fail-safe features and signal reliability have distinct advantages over Loran-C. Its excellent accuracy and immunity to atmospheric disturbances are obvious advantages over Omega. However, due to the long transitional process needed to introduce a new civil aviation NAVAID, the rapid replacement of the VORTAC system is not envisioned. Rather, a gradual, evolutionary transition to a GPS civil aviation navigation/collision avoidance/ communication/surveillance system is visualized, beginning in the early to mid 1980's.

6.0 TABLES

TABLE I

Qualitative rankings of specific navigation satellite system characteristics

NAVIGATION SYSTEM CHARACTERISTICS	MILITARY SIGNIFICANCE	CIVILIAN SIGNIFICANCE
Worldwide Coverage	Essential	Essential
Precision Accuracy	Essential	Desirable
Common Grid Capability	Essential	Essential
Passive non-transmitting Users	Essential	Desirable
Unlimited Number of Users	Essential	Desirable
Freedom from Ambiguities	Essential	Essential
High Anti-Jam Immunity	Essential	Desirable
Selective Denial to Un-authorized users	Essential	Minimal
Continuous Navigation Availability in real-time	Essential	Desirable
Operation with High Dynamic Users	Essential	Desirable
Minimal Frequency Allocation Needs	Essential	Desirable
All Weather Operation	Essential	Essential
Minimum Propagation Limitations	Desirable	Desirable
Satisfactory Form Factor, size, weight	Essential	Desirable
Minimal User Equipment Cost	Essential	Essential
Autonomous Ground Control from U.S. Territory	Desirable	Minimal
Compatibility with other navigation Systems	Desirable	Essential
Evolutionary transition into Operational Capability	Desirable	Essential
Provide world-wide Time and System Time Synchronization	Desirable	Desirable

TABLE II

Categorical error sources for GPS

Error Contributor	Category
Satellite Ephemeris Satellite Delay and Clock	Space Vehicle Segment
User Receiver Measurement Error User Mechanization Errors	User System Segment
Ionosphere Delay Tropospheric Delay Multipath	Propagation Link

TABLE III

GPS error budget, expressed as uncorrelated equivalent ranging errors

SOURCE	CONTRIBUTION, METERS, (1σ)	
	P-CODE	C/A-CODE
Satellite Ephemeris	1.52	1.52
Satellite Group Delay and Clock	0.91	0.91
Pseudorange Noise	1.0	10.0
Range Quantization	0.3	0.3
Mechanization Error	1.0	10.0
Ionospheric Dual Frequency Compensation	3.0	
Uncompensated Ionosphere		
Daytime		5.0 to 15.0
Nighttime		0.5 to 1.5
Troposphere	1.0	1.0
Multipath	1.2	5.0
RSS	4.0	16 to 21

TABLE V

Marine radiolocation systems

System Name	System Parameters				Function	Remarks
	Frequency	Coverage	Accuracy	Readout		
LORAN-A	1850, 1900, 2000 kHz	750-1500 mi. per chain	0.5-2 mi. (Repeatable) to 1000 mi.	3-5 min	Long-range navigation, oceanic and coastal	
LORAN-C	90-110 kHz	800-1200 mi. - groundwave; 3000 mi. - skywave	65-200 ft (Repeatable)	15 sec for manual, *continuous for automatic*	Long-range navigation, oceanic	
Omega	10.2, 11.3, 13.6 kHz	5000 mi. per station, *worldwide by 8 stations*	2.6-5.0 n.mi. by night, *2 n. mi. by day (Absolute)*	Position update every 2 *min to continuous*	Long-range navigation, worldwide	Lane ambiguity. Airborne and marine, commercial receivers available.
Differential Omega	10.2, 11.3, 13.6 kHz for OMEGA stations. Monitor station frequencies not determined	10-100 n.mi. from monitor station	0.5-1.2 n.mi. at 10-100 n.mi. (Absolute)	Continuous	Medium- to short-range navigation in areas where higher accuracy is required than Omega	Two receivers needed. In experimental stage.
ALPHA-Omega	OMEGA stations: 10.2, 11.3 and 13.6 kHz, ALPHA stations HF	100 n.mi.	500 ft (Repeatable)		Navigation to coastal and traffic confluence regions	In experimental stage
VLF Navigation (other than Omega)	USSHz 11.9, 12.0 and 14.8 kHz	Worldwide with a few stations (8 typical)	2.5-5.0 n.mi. (Absolute)	Once in 3-10 sec	Long-range navigation for ships and aircraft	
CONSOL	191 kHz	1000-1500 mi.	Azimuthal bearing within 0.7°. (Absolute)	Approximately 2 min	Line-of-position to transmitter	
DECCA	90-130 kHz	To 140 mi.	0.25-0.9 n.mi. (Absolute)	20 sec	Short- to medium-range coastal confluence	
Shipboard Radar	9300-9500 MHz 5160-5650 MHz 2900-3100 MHz	To 48 mi. usually 16 mi.	10 percent range 1° in azimuth (Absolute)	5 sec	Navigation and collision avoidance	Self-contained system
Shipboard Radar Beacons	9300-9500 MHz 5450-5650 MHz 2900-3100 MHz	To 48 mi.	10 percent, 1° in azimuth (Absolute)	5 sec	Navigation and collision avoidance	
TRANSIT Navy Navigation Satellite System (NNSS)	150 and 400 MHz	Worldwide	0.2 n.mi. (Absolute)	Varies with latitude from *110 min to 30 min*	Provides navigation fix and time	Utilizes 5 satellites in polar orbit

AD-A075 208 ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT--ETC F/G 17/7
PRINCIPLES AND OPERATIONAL ASPECTS OF PRECISION POSITION DETERM--ETC(U)
JUL 79 C T LEONDE
UNCLASSIFIED AGARD-AG-245 NL

4 OF 5
AD
A075208

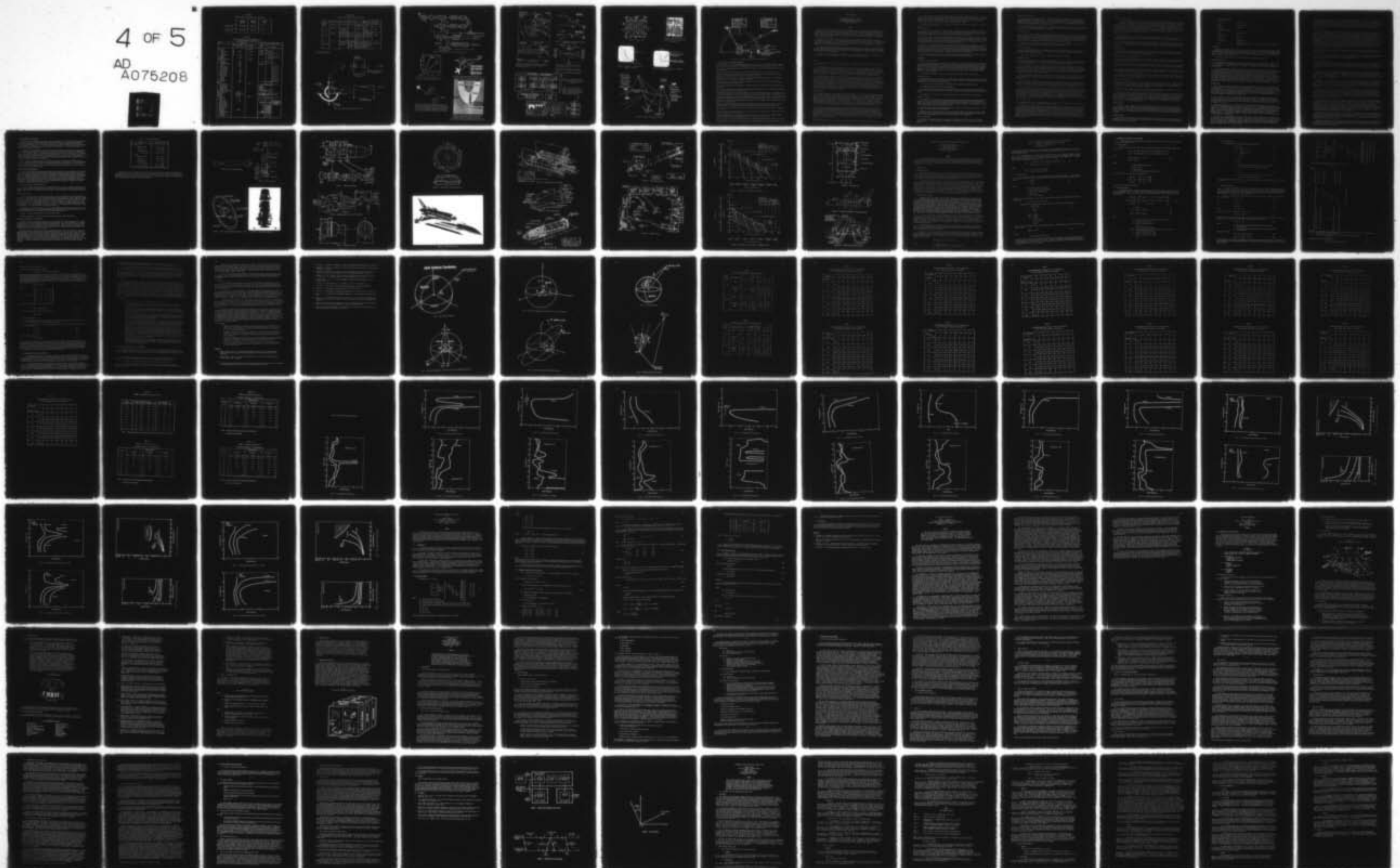


TABLE IV

Expected GPS civil user accuracy (C/A code)

	POSITION (METERS)	VERTICAL (METERS)	VELOCITY (M/S)
50% of time	20 to 30	30 to 40	0.2
90% of time	30 to 38	50 to 60	0.3

TABLE VI

Marine Navigation Accuracy Requirements
and User Population Distribution

User	Population		Area/Accuracy Requirement		
	Documented U.S. over 5 G.T.	World over 100 G.T.	High Seas	CCZ	HEZ
Merchant		34,347			
tanks					
tank barge	2,590	5,869	4NM	.25NM	
cargo					
cargo barge	17,956	27,708	4NM	.25NM	
lighter	33				
Passenger	6,805	770		.25NM	
Special Purpose		3,980			
Cable	10				
Dredging	494				6 ft
Oil Exploration	2,315		50 ft		
Ferry	267				
Fire Boats	39				
Ice Breaker	2				
pile driving	82				
pilot boat	104				
police boat	47				
patrol boat	53				
water boat	10				
whaler	2				
welding	12				
wrecking	22				
Tow boats	6,309			.25 NM	
Other		3,980			
hydrographic			.1 NM	8M-5M	
minerology			10 ft		
meteorology			1-3 NM		
bathymetry			3 NM		
special/scientific			0-0.1NM	0-0.1NM	
Miscellaneous	1,297				
Pleasure	1,400,000 est		0.5-1NM	0-0.1	
Yachting	37,924		0.5-1NM	0-0.1	
Fishing	21,583	11,949			
research				.5 NM	
COD	2				
Oystering	830				for safe nav
pots, netters			.025NM	100 ft	for safe nav
party boats					.025 NM
bottom travel				.025NM	
purse seine				0.2 NM	
drift gillnet				0.5 NM	
trolling	4,200			0.5 NM	
shrimping			0.5 NM	0.5 NM	
midwater trawl			0.5 NM	0.5 NM	
scalloper					
clammers					

TABLE VII

FAA RNAV and GPS Implementation Plans

RNAV Implementation Phase	FAA RNAV System Planning (per 1973) ..				NAVSTAR GPS System Capabilities	NAVSTAR Implementation Phase
	VOR Routes	2-D RNAV	3-D RNAV	Preplanned Routes		
I Pre-1977	Realign	Designated Enroute & Terminal	——	Limited Use	Test Program Phase	I 1977
II 1977-1982	Delete	2D RNAV is the Navigation System	Develop 3-D Procedures	Use where designated Routes are not available.	2D Capability Global Coverage	II 1980's
III Post-1982	——	Delete All Designated Routes	3D RNAV is The Navigation System	Basic route preplanned with random routing capability	Full GPS operational capability 2-D, 3-D and 4-D	III Post-1984

7.0 ILLUSTRATIONS

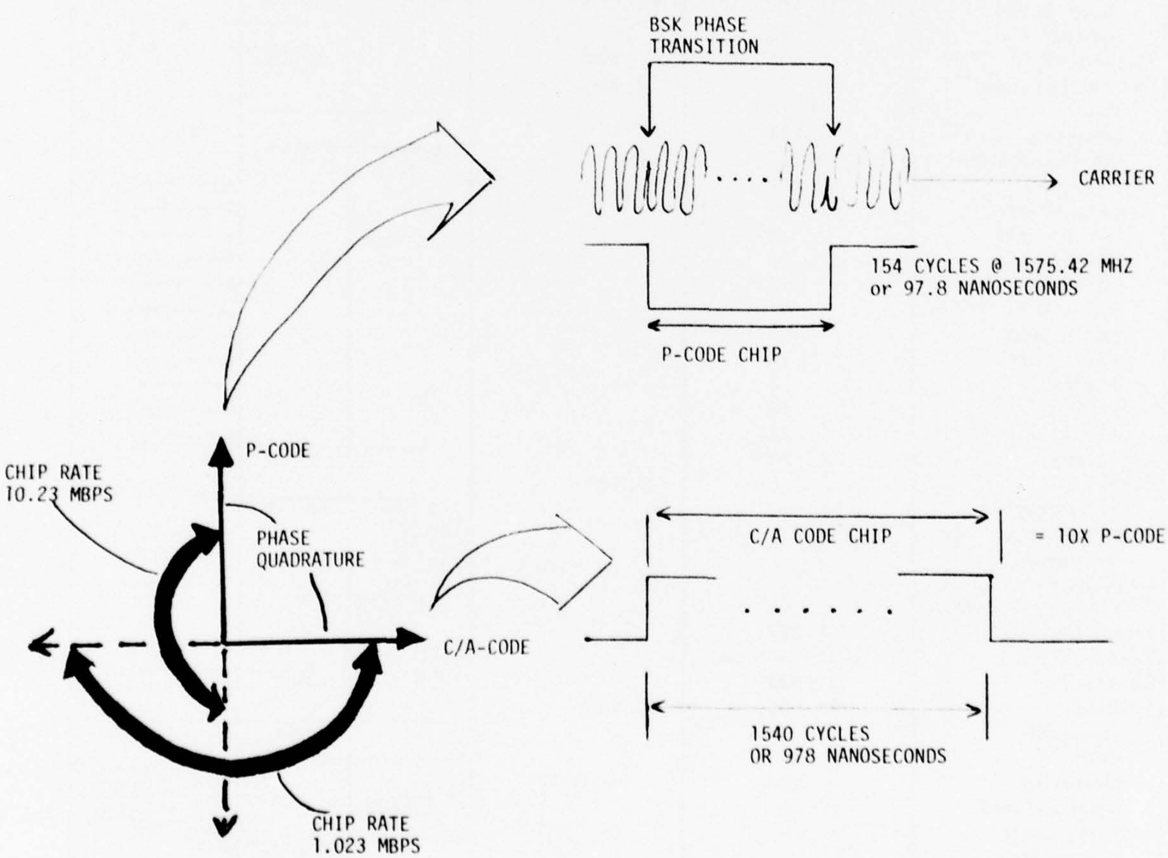


Figure 1 - GPS navigation signal structure.

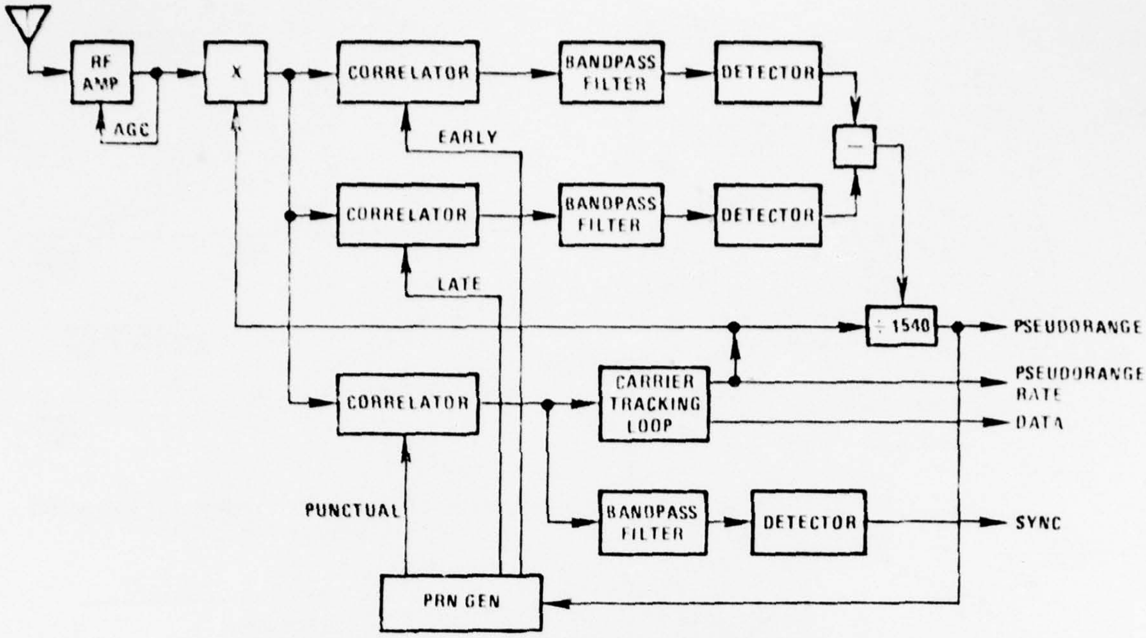


Figure 2 - Design for low cost civil GPS receiver

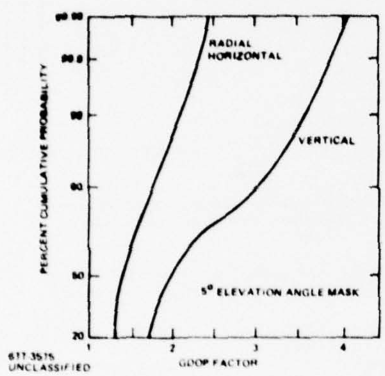


Figure 3 - Cumulative GDOP factor distribution for operational GPS phase.

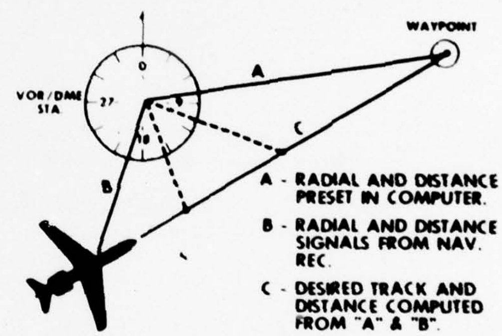


Figure 4 - Basic geometry for defining 2-D VOR/DME (VORTAC) RNAV waypoint in rho/theta

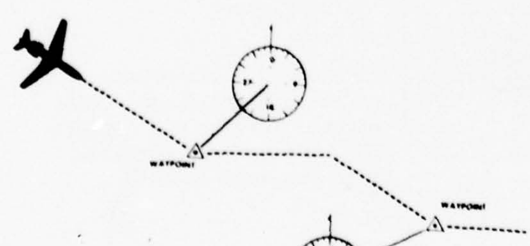


Figure 5 - Model of VOR/DME (VORTAC) route structuring. Tracking and distance measurement away from a rho/theta waypoint can be determined as well as to a waypoint. GPS route structuring could coincide but would be expressed in lat/long.

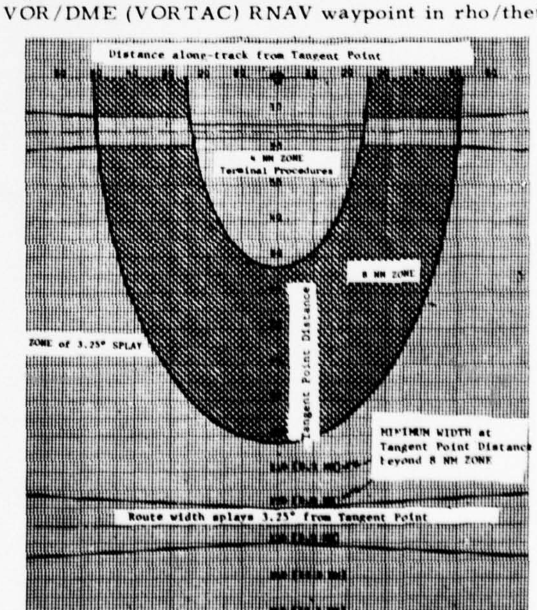


Figure 6 - AC 90-45A criteria for detemining VOR/DME (VORTAC) RNAV route widths. (These criteria form the basis at this time for FAA evaluation of all types of RNAV system accuracy.

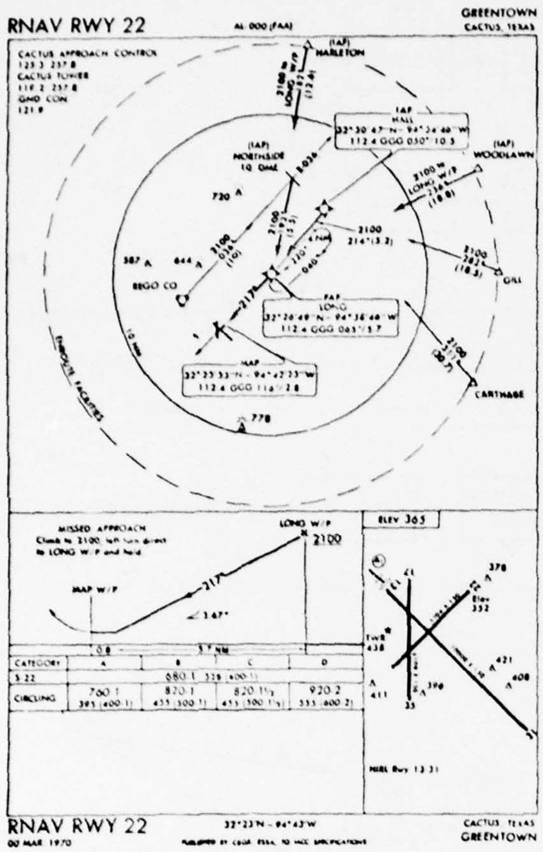


Figure 7 - Format of RNAV instrument approach chart (Both lat/lon and rho/theta waypoint coordinates are shown.)

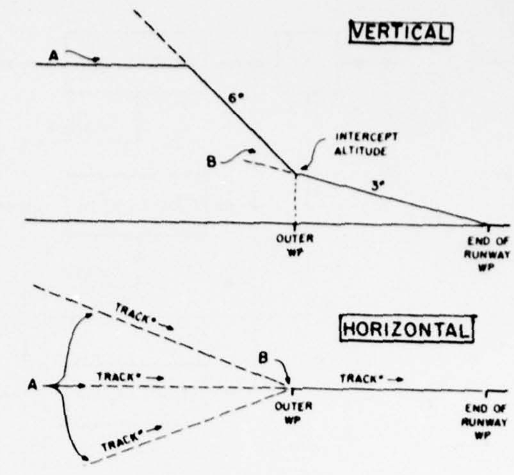
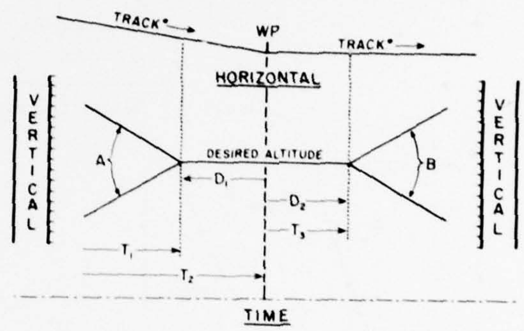


Figure 8 - Comparable vertical and horizontal profiles for 3-D RNAV.



A Descending/Ascending aircraft reaches desired altitude on desired track at desired distance (D_1) from Waypoint at desired time (T_1) and arrives at Waypoint at desired time (T_2)

B Ascending/Descending aircraft (on desired track) leaves desired altitude at desired distance (D_2) from Waypoint at desired time (T_3) for new desired altitude

Figure 9 - Four dimensional (4-D) matrix.

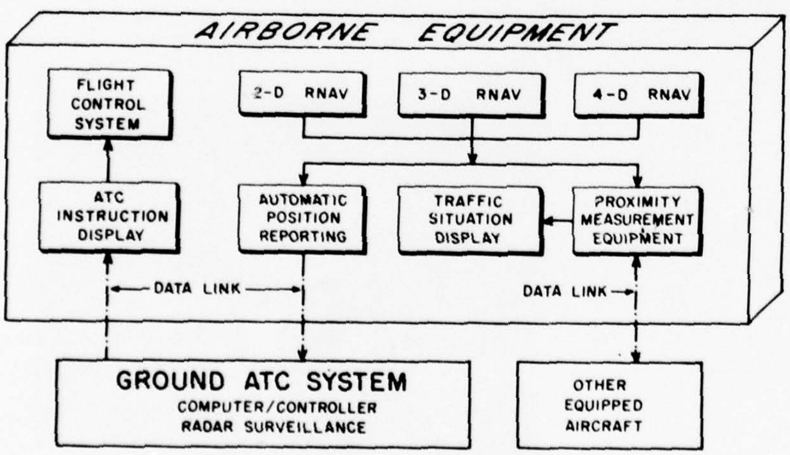
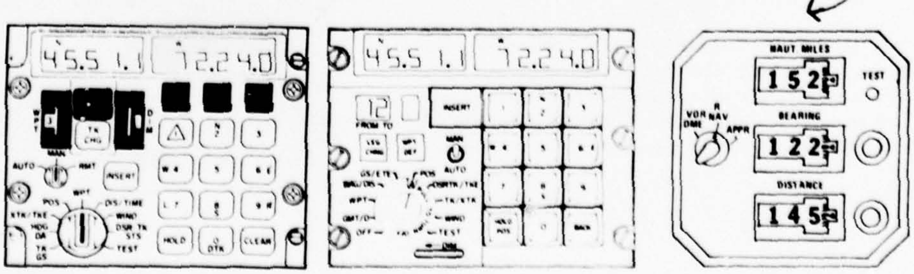


Figure 10 - Potential expansion of GPS capabilities in addition to navigation

Figure 11 - Representative RNAV control & display units (CDU). INS (A), VLF/OMEGA (B) and VOR/DME (VORTAC) single waypoint unit (C).



(A) (B) (C)

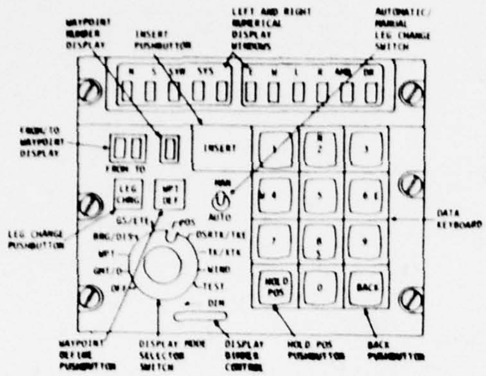


Figure 12 - Loran-C CDU

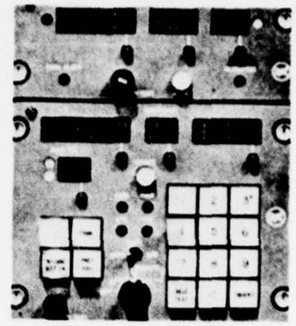


Figure 13 - VOR/DME (VORTAC) multiple waypoint CDU with 3-D programmer.

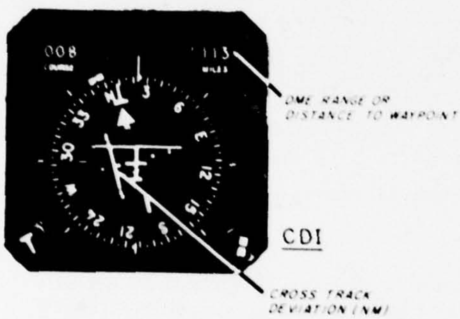


Figure 14 - Application of standard CDI/HSI to 2-D RNAV

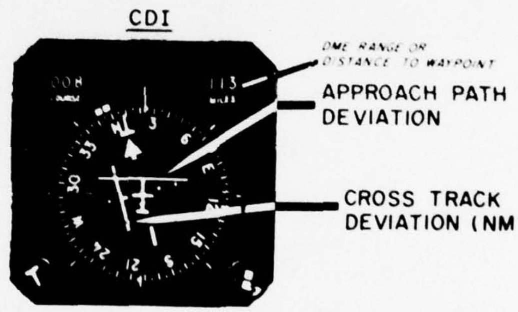


Figure 15 - Application of standard CDI/HSI to 3-D RNAV.

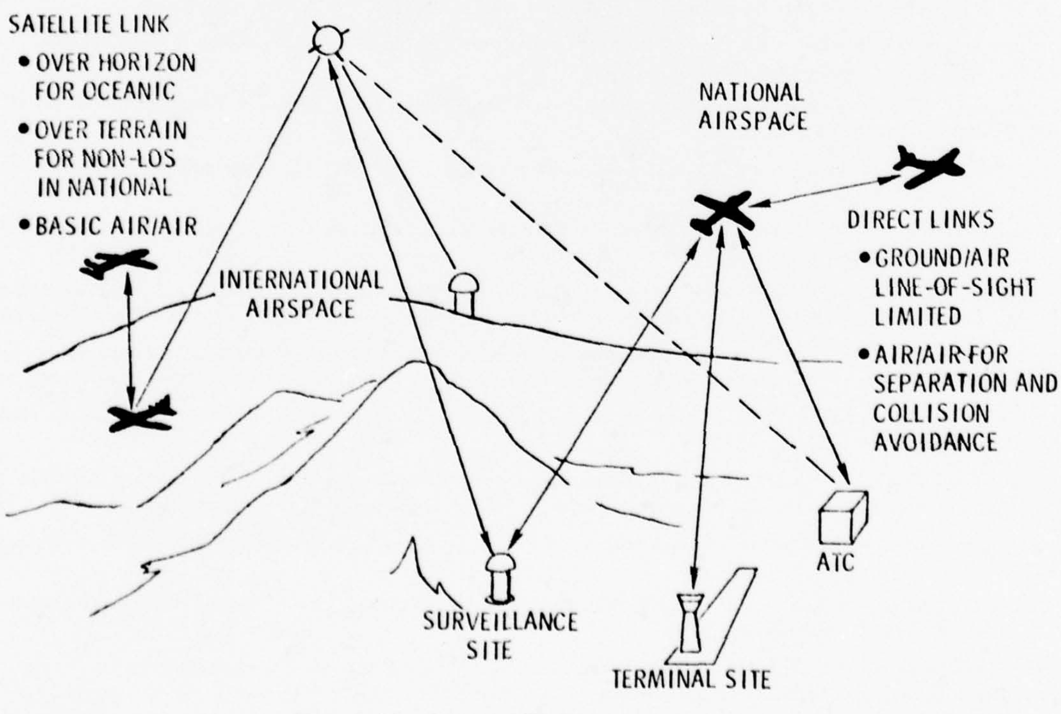


Figure 16 - Major GPS data link options.

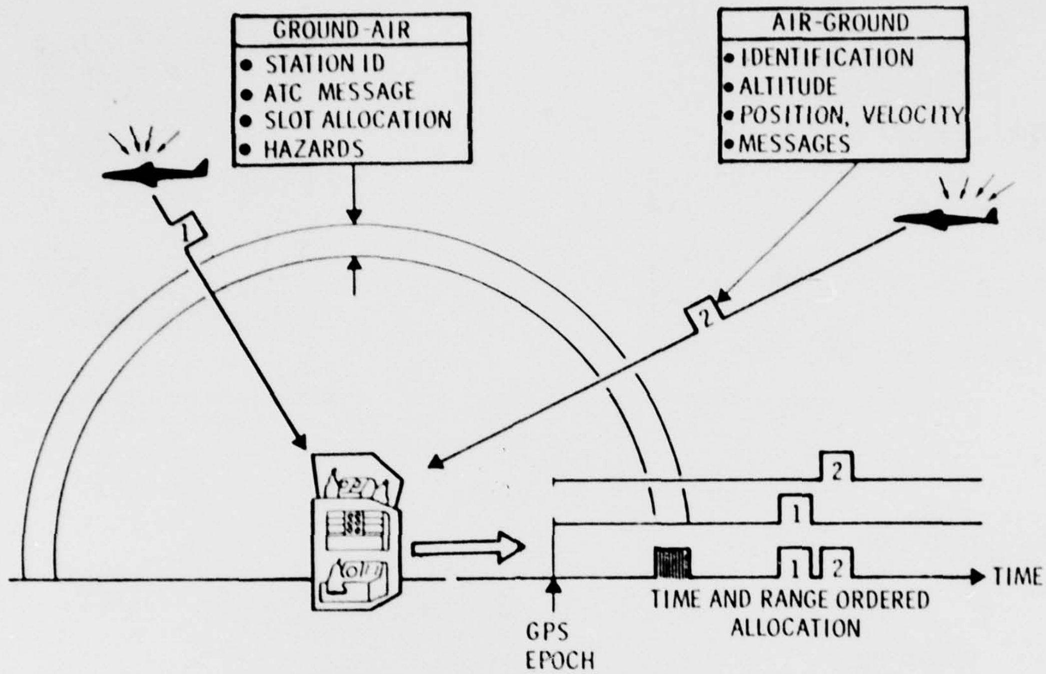


Figure 17 - Communication data link and identification.

8.0 REFERENCES AND BIBLIOGRAPHY

K. D. McDonald, "A Survey of Satellite-Based Systems for Navigation, Position Surveillance, Traffic Control, and Collision Avoidance", Navigation, Journal of the Institute of Navigation, Vol 20 1973-74.

E. H. Martin, "GPS User Equipment Error Models", Navigation, Journal of the Institute of Navigation, Vol 25, Number 2, 1978.

"Automotive Vehicle Locator Systems in Land Mobile Radio Services", FCC Docket No. 18302.

"A Guide for Determining the Cost Effectiveness of Automatic Vehicle Monitoring Systems in Urban Applications", Integrated Traffic Systems, Inc. ITS701, July 1970.

W. S. Murray and W. C. Scales, "Overview of Automatic Vehicle Monitoring Systems", The MITRE Corporation, MTP-386, McLean, Virginia, August 1973.

S. Riter, Editor, "Proceedings of an Urban Technology Conference on Automatic Vehicle Monitoring" Texas A&M University, November 1972.

S. H. Roth, "Automatic Vehicle Monitoring Technology Review", The MITRE Corporation, MTR-6059, August 1971.

W. C. Scales, "Benefits and Costs of an Automatic Vehicle Monitoring System", The MITRE Corporation, MTR-6064, McLean, Virginia, September 1971.

"Radio Navigation Systems Economic and Planning Analysis" Final Report, Volume 2, Computer Sciences Corporation, Prepared for Office of Telecommunications Policy, Executive Office of the President, 1977.

"Radio Navigation Systems Economic and Planning Analysis" Final Report, Volume 2, Computer Sciences Corporation, Prepared for Office of Telecommunications Policy, Executive Office of the President, 1977.

"Department of Transportation National Plan for Navigation, U. S. Department of Transportation, Washington, D. C., November 1977.

"Preliminary Study of NAVSTAR/GPS for General Aviation", Research Triangle Institute, 430-1228, NASA CR-145059, Final Report, Prepared for NASA Langley Research Center, November 1976.

Glen A. Gilbert "Air Traffic Control: The Uncrowded Sky", Smithsonian Institution Press, Washington, D. C., 1973

"Status of the NAVSTAR Global Positioning System", Report to the Congress by the Comptroller General of the U. S., PSAD-77-23, March 1977.

"Global Positioning System", Special Studies Task C1063D, Magnavox, Advanced Products Division, Torrance, California , for Joint Program Office, U. S. Air Force, March 1976.

Glen A. Gilbert, "ATC System Development Trends", Flight Operations Magazine, Dallas, Texas, 1977

"Navigation Planning - Need for a New Direction", Comptroller General Report to the Congress, 1978

LAUNCH VEHICLES

William Nissim and Curt J. Zoller
Advanced Program Department
Rockwell International Space Division
Downey, California

SUMMARY

The GPS satellites will be deployed to their operational orbit by means of the Atlas F launch vehicle (manufactured by General Dynamics/Convair Division), which carries the satellite and its upper stage to a suborbital trajectory, and by an upper stage (manufactured by Fairchild Space and Electronics Co.), which provides the transfer capability from the suborbital trajectory to an apogee altitude of 10,898 nautical miles. At apogee the apogee kick motor in the satellite ignites and injects the satellite into a 10,898-nautical mile circular orbit.

Operational GPS satellites will be deployed from the Space Shuttle (under development by Rockwell International), which has the capability of deploying several GPS satellites per flight. The stage vehicles utilized with these missions will be either the inertial upper stage (IUS), one of which can carry up to four GPS satellites; individual propulsive stages such as the spinning solid upper stage (SSUS); or a special stage vehicle such as used in the Atlas F launches.

The launch vehicles and upper stages are described in this section. The description includes physical and performance characteristics and the expected operational procedure, based on data obtained from various Rockwell International publications (for Shuttle), and from representatives of General Dynamics (for Atlas) and Fairchild (for the stage vehicle).[1]

Potential system improvements being considered for the operational phase will result in satellite weight growth, which can be accommodated by launch and stage vehicle growth options.

ATLAS F

The launch vehicle for the Phase 1 GPS program is an F-series Atlas booster, modified by a mission-peculiar kit.

The fundamental Atlas design characteristic is the one and a half stage concept, in which all three engines are ignited to initiate launch and all are fed from the same fuel and oxidizer tanks. Two of the engines are mounted on the half stage (booster thrust section). The third engine and the propellant tanks comprise the sustainer section. The booster thrust section is staged after the vehicle weight has decreased (due to utilization of propellants) to a value which allows the sustainer engine to maintain a desired acceleration (Figure 1). Staging time can be varied with changes in trajectory and payload weight.

The Atlas booster airframe is of stainless steel and aluminum structure 3 meters (10 feet) in diameter and more than 21.3 meters (70 feet) long. Figure 2 depicts its principal design features. Its empty weight is about seven tons and it is composed of three sections: the payload adapter, the sustainer section, and the booster thrust section. The adapter connects the upper stage vehicle to the sustainer section. The tank section is a stainless steel monocoque shell providing tank space for approximately 11 ,500 liters (30,000 gallons) of propellants and structural support for the payload vehicle, as well as mounting points for all major subsystem components. The booster section consists of an aluminum cylinder constructed belt frames, longerons, and a reinforced skin, with nacelles and fireshield of aluminum-fiberglass honeycomb. The Atlas airframe provides an optimum design compromise between low weight and high strength which has proven itself to be dependable and effective.

Sustainer Section

The sustainer section is cold-rolled stainless steel formed into a 3-meter (10-foot) diameter cylinder tapering at the forward (payload adapter) end, with bulkheads at both end. The bottom connects with the booster section. The tank is divided into two sections, separated by an intermediate bulkhead. The upper portion of the tank section contains approximately 77,500 kilograms (171,000 pounds) of liquid oxygen (LO_2). The lower portion of the fuel tank contains approximately 34.500 kilograms (76,000 pounds) of fuel. The tank body ranges in thickness from 5 cm (0.038 inch) to . 0. cm (0.011 inch). The main or sustainer engine is located on the centerline at the aft end of the vehicle. Mounted along the sides of the propellant tanks are two small vernier motors. The sustainer engine is gimbal-mounted and provides directional control. After this engine is shut down, the verniers allow correction in pitch, yaw and roll.

The vehicle tanks are pressurized at all times because the tank section is not a rigid structure. (Some subsystem checkout must be accomplished with the tanks vented and the vehicle in stretch.) The fuel tank contains more pressure than the LO_2 tank in order to keep the intermediate bulkhead from reversing itself. Normally, the vehicle is kept in stretch by use of a stretch mechanism which is attached to the vehicle and the tower. A load of 44,500 newtons (10,000 pounds) is put on the vehicle in order to stretch and maintain a rigid structure. The stretch is used as a backup in case of a pressure loss. The pressure-sensing probes in the tanks control the pressure regulators in the booster section to maintain the proper pressures in the propellant tanks.

---

[1] Atlas F data were provided by Mr. Phil Genser of General Dynamics/Convair; stage vehicle data were supplied by Mr. Joe Hendricks of Fairchild Space and Electronics Co.

The sustainer engine is mounted at the centerline on the aft end of the sustainer structure. It has a thrust rating of 253,548 newtons (57,000 pounds). Two vernier engines rated at 8,896 newtons (2000 pounds) thrust each are mounted on opposite sides of the tank section in the yaw plane and provide roll control during sustainer engine operation and velocity trim after sustainer engine cutoff. The sustainer and vernier engines provide thrust during the entire powered portion of flight.

Velocity control and stabilization are achieved by gimbaling the engine thrust chamber to control the direction of the thrust vector.

Booster Section

The booster section consists of a thrust cylinder, a forward nacelle, an aft nacelle, and a booster radiation shield. The booster section houses the two booster engines and the one sustainer engine and protects them from radiant heat. The booster section also provides support and connection for the various engine components and accessories. At staging, the booster section is jettisoned.

The thrust cylinder is a reinforced aluminum shell 3 meters (10 feet) in diameter and 1.8 meter (6 feet) long, with a 1.5-meter (5-foot) extension covering the area between the missile support longerons. Longitudinal stiffeners are attached to the exterior of the shell for additional strength and give the cylinder a corrugated appearance. The forward end of the thrust cylinder connects to the main thrust ring on the fuel tank and the aft edge connects to the booster radiation shield.

The purpose of the thrust cylinder is to transmit the thrust from the two booster engines to the main thrust ring. It also supports the subassemblies of the booster stage propulsion, pneumatic, electrical, and hudraulic systems. Just aft of each vernier engine located on the exterior of the thrust cylinder are two vernier engine radiation shields. The shields are made of laminated fiberglass and protect the booster section from the vernier engine exhaust. The nacelles are constructed of aluminum honeycomb core bonded to laminated fiberglass skins. The nacelles extend the length of the booster section.

The booster engines are rated at 165,000 pounds of thrust each, thus giving a total thrust capability at liftoff of 330,000 pounds plus 57,000 pounds of thrust from the sustainer engine.

Subsystems

Propellant Utilization Subsystem

The propellant utilization subsystem provides a means of controlling propellant flow with more accuracy than the fixed metering orifices in the propellant feed system are capable of maintaining. Such control eliminates any unbalanced utilization of propellants by the engines which might otherwise limit the performance of the vehicle. Propellant consumption is, therefore, scheduled to result in a minimum of residual propellants. The scheduling is accomplished by a propellant utilization valve installed in the sustainer engine fuel line. The system is capable of causing a change of $\pm$ 15 percent in the propellant mixture ratio for the sustainer engine.

The propellant utilization subsystem is the Acoustica liquid-level sensor control system.

Pneumatic Subsystem

During launch and flight, the pneumatic subsystem supplies helium at regulated pressure for pressurization and control functions of the boost vehicle. This subsystem maintains the required pressure levels and differential for the LO_2 and fuel tanks to ensure:

1. Vehicle tank structural rigidity and integrity under inertial and aerodynamic loads.

2. Adequate pressure head at propellant turbopumps.

The pneumatic subsystem also supplies required pressurization to turbopump lube tanks and hydraulic reservoirs, for booster stage separation latch release, and for operation of main and vernier engine control units.

Six spherical bottles inside the thrust section supply the helium required for pressurization.

Hydraulic Subsystem

The space launch vehicle contains two independent hydraulic systems supplying the operating pressure required to position the engine thrust chambers for directional and roll control of the vehicle during flight. One system is for the booster engines; the other is for the sustainer and vernier engines. Electrical signals from the autopilot are applied to hydraulic actuator assemblies, each consisting of a servo valve, an actuating cylinder, and a linear transducer.

Hydraulic power for the booster engine hydraulic system comes from a hydraulic pump driven by a power takeoff on the booster-engine turbopump assembly.

Hydraulic power for the sustainer system is obtained from a hydraulic pump driven by a power takeoff on the sustainer engine turbopump assembly until sustainer engine cutoff. After sustainer engine cutoff, power for the vernier-engine actuators on is supplied by vernier, solo accumulators.

Electrical Subsystem

The airborne electrical subsystem is composed of a 28-vdc vehicle battery and a 115 vac, 3-phase, 400-Hz inverter. A change-over switch is provided for switching both ac and dc power from external to internal.

Flight Control Subsystem

The flight control subsystem consists of an autopilot system (including a flight programmer) and 10 gimbaled thrust-chamber actuator assemblies. The subsystem stabilizes and steers the vehicle along a desired flight path by controlled orientation of the engine thrust vectors. Pitch steering commands are generated by the flight programmer from launch until staging and by the flight programmer and the guidance subsystem from staging until the end of powered flight.

The autopilot system is custom-tailored for each individual space mission. The mission-peculiar variations in the autopilot kit are primarily in the timing and sequencing switches and related circuitry. A new Convair high reliability autopilot is now in use.

Guidance Subsystem

The GE radio inertial guidance subsystem determines the position and velocity of the vehicle, and, on the basis of the trajectory objectives, commands the flight correction necessary to fulfill these objectives. A ground-based monopulse, X-band tracking radar transmits composite bursts which contain identity information, interrogates an airborne position transponder (pulse beacon), and commands discrete functions and attitude changes. Position and velocity information flow to data extraction and processing circuits in the ground computer, where commands are generated to accomplish trajectory objectives.

Information from the autopilot controls gimbaling of the engine thrust chambers in such a way as to guide the vehicle onto the desired flight path, thus closing the guidance loop.

Telemetering Subsystem

The telemetering subsystem is the instrumentation required to transform physical variables such as temperature and acceleration into electrical signals which are transmitted to ground for the purpose of securing technical data during flight.

A range safety command destruct system is available for emergency situations.

Deployment of GPS

Phase I GPS satellites will be launched from Vandenberg Air Force Base (VAFB) using an Atlas-F/stage vehicle launch system (Figure 3). The orbital period of each satellite will be nearly 1/2 sideral day and thus the ground trace of each of the satellites will repeat on a day-to-day basis.

At the termination of boost by the launch vehicle, the satellite will be separated from the launch vehicle's final stage in a spin-stabilized mode and placed in an elliptical transfer orbit. While in the transfer orbit, the vehicle will be commanded by ground control to perform the necessary orbital maneuvers in preparation for insertion into an initial drift orbit. These maneuvers include an approximately 180-degree satellite attitude change (to align the apogee kick motor thrust vector properly), and minor attitude trim to compensate for control system.

After the satellite has been properly oriented, ignition of the apogee kick motor (AKM) will be commanded by Satellite Control Facility (SCF). This will be accomplished when the satellite is near apogee in the transfer orbit. At apogee kick motor burnout, the satellite will be in near-circular initial drift orbit.

While in the initial drift orbit, the satellite will be commanded by the SCF to perform delta velocity maneuvers if required to correct for orbit period and eccentricity errors introduced by the apogee kick motor burn. The need for the maneuvers will be determined based on the satellite ephemeris as determined by the SCF. At the conclusion of the maneuvers, the satellite will be in a final drift orbit.

In the final drift orbit, the satellite will be despun and stabilized in three axes through sequences of commands from the SCF. The satellite will first be despun to a low rate. Acquisition of the earth by the satellite control function will then be accomplished, followed by deployment of one of the solar array wings. Acquisition of the sun will be accomplished while the satellite is in that configuration. The second solar array wing then will be deployed and the satellite subsystems configured to operate in an autonomous mode. A series of delta velocity maneuvers will be commanded by SCF to position the vehicle in its final orbit. The system deployment mission phase concludes with the final delta velocity maneuver that places the satellite in its proper position in the GPS Phase I constellation.

FAIRCHILD UPPER STAGE

The upper stage utilized in conjunction with Atlas F launches is manufactured by the Fairchild Space and Electronics Co. A picture of the stage is shown in Figure 4.

The 3.1-meter (10.2-feet) long consists of three principal parts, called Stage 0, Stage 1, and Stage 2. Figure 5 depicts the various components of the stage, dimensional data, and the installation of the stage on the Atlas F launch vehicle. Stage 0 consists of the Atlas interface structure, the destruct system, separation system, and electrical components and wiring. Stage 1 includes a solid rocket motor (PKM No. 1), the stabilization system, a motor interstage structure, an aft motor structure, separation system, and electrical components and wiring. Stage 2 is composed of the spacecraft interface structure, a solid rocket motor (PKM No. 2), telemetry, forward motor structure, separation system, and electrical components and wiring. Interconnections between the various devices in each stage and between the three stages are illustrated in Figure 6.

Perigee Kick Motors (PKM's)

The PKM's are manufactured by the Thiokol Chemical Corp. Each motor, Model TE-M-364-4, is a 0.96-meter (38-inch) diameter, 1.67-meter (66-inch) long solid propellant motor with a 6 Al-4V titanium case and loaded with 1038 kilograms (2290 pounds) of TP-H-3062 Class 2 case bonded composite solid propellant. The TE-M-364-4 motor is shown in Figure 7. Motor ignition is accomplished with a small solid-propellant pyrogen discharging its gases into the main motor cavity. The ignition chain is started by electrically firing redundant squibs. A remote electromechanical safe and arm (S/A) device (Model TE-O-642-4) prevents premature motor ignition in case of accidental squib ignition. The squibs are located in the remote S/A device. The ignition impulse is transferred from the S/A device to the PKM through-bulkhead initiator by the use of redundant interconnect trains (mild detonating cords).

Stabilization System

The stabilization system provides the capability to spin the stage vehicle and space vehicle to approximately 95 RPM. It is a hot gas generator system that provides thrust from six generators through 12 lines to six pairs of nozzles.

The thrust is generated by six ARC/Marc 67A2 hot gas generators. Each generator is a cylinder 7.62 by 20.3 cm (3 by 8.6 inches) which weighs 1.31 kilograms (2.89 pounds) and uses 0.77 kilograms (1.7 pounds) of ARCITE 479 Class 2 solid propellant. The ignition chain is started by electrically firing redundant initiators (squibs).

Structural System

The stage vehicle structural system (see Figure 5) is comprised of five major segments: the Atlas interface structure (which mates with the Atlas conical adapter), the aft motor structure, the forward motor structure, and the space vehicle interface structure.

Electrical Power and Distribution System

The electrical power and distribution system provides control and distribution of electrical power to the stage vehicle equipment. The subsystem consists of three separate sections, independent of the others and each featuring total redundancy.

The Stage 0 electrical power system provides the functions for stage vehicle-launch vehicle separation. It also contains the capability to initiate destruction of PKM 1 and PKM 2 in the event of missile malfunction.

The Stage 1 electrical system provides the timing, power, and distribution for the events which follow stage vehicle-launch vehicle separation. These events are: sequential ignition of the spin system, ignition of PKM 1, and Stage 1-Stage 2 separation.

The Stage 2 electrical system provides the remaining equipment timing, power, and distribution functions that inject the space vehicle into transfer orbit. The functional events are ignition of PKM 2 and separation of the space vehicle from the stage vehicle.

Ordnance and Separation System

Separation is achieved by explosive clamp separator bolts, Hi-Shear SC 1006-6 series. There are three clamp separators in the Atlas-stage vehicle separation Marman clamp arrangement and two clamp separators in the stage vehicle's Stage 1-Stage 2 and Stage 2-space vehicle clamp arrangements. Each clamp separator contains a single Hi-Shear PC 60 series power cartridge.

Initiation of the hot gas generator spinup system is achieved by redundant Space Ordnance Systems, Inc., SBASI series initiators. There are two initiators per generator (12 total).

Ignition of the PKM's is achieved by initiators in the remotely mounted PKM safe and arm devices. There are two initiators, Space Ordnance Systems SBASI-type series, in each safe and arm device. These initiators ignite redundant fuses (mild detonating cords) which detonate a charge of the through-bulkhead initiator mounted on the PKM. Detonation of this charge causes a high-order pressure wave which is transmitted through a steel diaphram to the receptor charge portion of the initiator (the steel diaphram is not physically penetrated). The receptor charge is detonated which in turn ignites the PKM pyrogen igniter, which discharges its gases into the main motor cavity.

Destruct System

The destruct system, located in the aft interface structure provides for destruction of the stage vehicle PKM's by a shaped charge that penetrates both solid rocket motor cases in the event of a malfunction prior to separation from the Atlas booster, a premature separation from the Atlas, or a premature Stage 1-Stage 2 separation. Destruction is initiated by an uplink command signal to the Atlas or by onboard switch actuation in the event of premature separation of the stage vehicle from the Atlas booster of of Stage 1 and Stage 2.

Telemetry System

The telemetry system provides in-flight performance data so that anomalies, if any, may be observed and measures taken to prevent recurrence on future flights. The characteristics of the telemetry transmitter and telemetry antenna are as follows:

1. Telemetry Transmitter

Frequency:	2252.5 MHz
Power:	5 watts minimum
Modulation:	PAM/FM/FM

2. Telemetry Antenna

Type:	Printed circuit
Polarization:	Linear
Location:	Station 373
Orientation:	Omnidirectional
Gain:	-10dB minimum with respect to isotropic over 90% of the radiation sphere.

SPACE SHUTTLE

Beginning in the 1980's routine access to space will be provided by the Space Shuttle System, a principal element of the Space Transportation System which also includes the Inertial Upper Stage (IUS) among its other elements. The Space Shuttle flight system (Figure 9) consists of the orbiter, an external tank that contains the ascent propellants to be used by the orbiter main engines, and two solid rocket boosters (SRB's). Figure 10 gives dimensional data.

The orbiter and SRB's are reusable; the external tank is expended in each launch.

Space Shuttle Orbiter

The orbiter spacecraft (Figure 11) contains the crew and payload for the Space Shuttle system. The orbiter can deliver to orbit payloads of 29,500 kilograms (65,000 pounds) with lengths to 18 meters (60 feet) and diameters of 5 meters (15 feet). The orbiter is comparable in size and weight to modern transport aircraft.

The three main propulsion rocket engines used during launch are contained in the aft fuselage. The rocket engine propellant is contained in the external tank, which is jettisoned before initial orbit insertion. The orbital maneuvering subsystem (OMS) is contained in two external pods on the aft fuselage. These units provide thrust for orbit insertion, orbit change, rendezvous, and return to earth. The reaction control subsystem (RCS) is contained in the two OMS pods and in a module in the nose section of the forward fuselage. These units provide attitude control in space and precision velocity changes for the final phases of rendezvous and docking or orbit modification. In addition, the RCS, in conjunction with the orbiter aerodynamic control surfaces, provides attitude control during reentry. They take effect in the lower, more dense atmosphere to provide control of the orbiter at speeds less than Mach 5. The orbiter is designed to land at a speed of 95 m/sec (185 knots), similar to current high-performance aircraft.

External Tank

The external tank (Figure 12) supplies the orbiter main propulsion system with liquid hydrogen (LH_2) and liquid oxygen (LO_2) at prescribed pressures, temperature, and flow rates. Both the LH_2 and LO_2 tanks are equipped with a vent and relief valve to permit loading, pressurization, and relief functions. Tank level sensors provide for propellant loading and shutdown signals. The external tank is thermally protected with a nominal 1-inch thick spray-on foam insulation (SOFI); extra insulation and a charring ablator (SLA 561) are used to withstand localized high heating. Since the tank is an expendable element, its subsystems are designed for single usage to minimize costs.

The external tank reacts the solid rocket booster thrust through its intertank and provides attach structure to the orbiter to react the Space Shuttle main engine thrust. At liftoff, the tank contains 703,000 kilograms (1.55 million) pounds of usable propellant. At main engine cutoff, it is separated from the orbiter before orbital velocity is achieved. The tank then proceeds on a ballistic reentry path for a safe impact in the ocean.

The external tank consists of a forward LO_2 tank, an unpressurized intertank, and an LH_2 tank. The LO_2 tank, with a volume of 551.85 cubic meters (19,500 cubic feet), is an aluminum alloy monocoque structure composed of a fusion-welded assembly of preformed, chem-milled gores, panels, machined fittings, and ring chords. The LO_2 tank is designed to operate at a nominal pressure range of 20 to 22 psig. The tank contains antislosh and antivortex baffles as well as an auto-geyser system to minimize liquid residuals and to damp fluid motion. A 43-cm (17-inch) diameter feedline conveys propellant through the intertank and externally aft to the tank-orbiter disconnect. The tank's double wedge nose cone reduces drag and heating and also serves as a lightning rod.

The intertank is a semimonocoque cylindrical structure with flanges on each end for joining the LO_2 and LH_2 tanks. It contains the SRB thrust beam and fittings which distribute SRB loads to LO_2 and LH_2 tanks. It houses instrumentation components and provides an umbilical plate which interfaces to a ground facility arm. The umbilical plate accommodates purge gas supply, hazardous gas detection, and hydrogen gas boiloff during ground operations. The intertank consists of mechanically joined skin, stringers, and machined panels of aluminum alloy and is vented in flight.

The LH_2 tank, with a volume of 1,573. cubic meters (55,552 feet^3), is a semimonocoque structure composed of fusion-welded barrel sections, five beam ring frames, and forward and aft 0.75 ellipsoidal domes. The LH_2 tank is designed to operate at a nominal pressure of 32 to 34 psia. It contains an anti-vortex baffle and a siphon outlet to transmit propellant to the tank-orbiter disconnect through a 43.2-cm (17-inch) diameter line. The LH_2 tank has provisions for the tank-orbiter forward attach strut, two tank-orbiter aft attach fittings, the thrust distribution structure, and the aft SRB-tank stabilizing strut attachments.

Solid Rocket Booster

Two solid-rocket boosters burn in parallel with the orbiter main propulsion system to provide initial ascent thrust. Primary elements of the booster (Figure 13) are the motor, including case, propellant, igniter, and nozzle; structural systems; separation, operational flight instrumentation, and recovery avionics; separation motors and pyrotechnics; deceleration system; range safety destruct system; and thrust vector control subsystem. Each SRB weighs approximately 586,500 kilograms (1.293 million pounds) and produces 2.90 million pounds of thrust at sea level. The propellant grain is shaped to reduce thrust approximately one-third 55 seconds after liftoff to prevent overstressing the vehicle during the period of maximum dynamic pressure. The grain is of conventional design, employing a star perforation in the forward motor closure and a double truncated cone perforation in each of the segments and aft closure. The contoured nozzle expansion ratio (area of exit to area of throat) is 7.16. The SRB thrust vector control is a closed-loop hydraulic system with power provided by redundant APU's and hydraulic pumps, has an omni-axial gimbal capability of 7.1 degrees which, in conjunction with the orbiter main engines, provides the flight control during the Shuttle boost phase.

The SRB's are released by pyrotechnic separation devices at the forward thrust attachment and the aft sway braces. Eight separation rockets on each SRB, four aft and four forward, separate the SRB from the orbiter and tank.

The forward section provides installation volume for the SRB electronics, recovery gear, range safety destruct system, and forward separation rockets. The parachute deceleration subsystem consists of a pilot parachute, a ribbon drogue parachute, and ribbon main parachutes.

Space Shuttle Mission

The Space Shuttle system will reduce the costs of earth-orbital operations while improving operational capabilities and flexibility. With a due-east launch from Kennedy Space Center (KSC), the Shuttle can deliver payloads up to 65,000 pounds to a 150-nautical-mile circular orbit. In addition, the Space Shuttle can return to earth with up to 32,000 pounds of payload, a capability not provided by expendable launch vehicles.

The Shuttle is launched with the three orbiter Space Shuttle main engines and the two solid rocket boosters burning in parallel (Figure 14). A maximum dynamic pressure (Q) of 650 pounds per square foot is experienced 62.5 seconds after launch at an altitude of 37,100 feet. At 108 seconds, the X-axis load factor reaches a maximum value of 2.5 g's. SRB separation occurs at 122.3 seconds at an altitude of 142,200 feet, 25 nautical miles downrange from the launch site. After SRB separation, the orbiter continues to ascent, using the three main engines.

Main engine cutoff takes place 479 seconds after liftoff, when the orbiter has reached an altitude of 379.740 feet. The external tank separates at main engine cutoff. The OMS engines provide the additional velocity needed to insert the orbiter into an elliptical orbit having a minimum apogee of 150 nautical miles. The OMS engine cutoff occurs 600 seconds after launch at an altitude of 418,650 feet, when the orbiter is 1428.2 nautical miles from the launch site. At first apogee, the orbiter initiates the first of two maneuvers to circularize the orbit at 150 nautical miles.

Following the completion of orbital operations, the orbiter is oriented to a tail-first attitude. After the OMS provides the deceleration thrust necessary for deorbiting, the orbiter is reoriented nose-forward to the proper attitude for entry. The orientation of the orbiter is established and maintained by the reaction control system down to the attitude where the atmospheric density is sufficient for the pitch and roll aerodynamic control surfaces to be effective (about 250,000 feet altitude and 26,000 feet per second velocity). The yaw RCS remains active until a vehicle reaches an angle of attack of about 10 degrees (about 80,000 feet altitude).

The orbiter entry trajectory provides lateral flight range to the landing site and energy management for an unpowered landing. The trajectory, lateral range, and heating are controlled through the attitude of the vehicle by angle of attack and bank angle. The angle of attack is established at 38 degrees for the theoretical entry interface of 8,500 feet altitude. The entry flight path angle is -1.19 degrees. The 38-degree attitude is held until the speed is reduced to 21,200 feet per second (about 220,000 feet altitude), is then reduced gradually to 28 degrees at 17,200 feet per second (about 190,000 feet altitude); it is held at 28 degrees until speed is reduced to 8500 feet per second (about 150,000 feet altitude), and then reduced gradually to 6 degrees where the speed is about 1500 feet per second (about 70,000 feet altitude) at the beginning of terminal area energy management.

During the final phases of descent, flight path control is maintained by using the aerodynamic surfaces. Terminal area energy management is initiated to provide the proper vehicle approach to the runway with respect to position, energy, and heading. Final touchdown occurs at an angle of attack of about 16 degrees. The maximum landing speed for a 14,500-kilogram (32,000-pound) payload, including dispersions for hot day effects and tailwinds, is about 207 knots.

The orbiter normally carries into orbit a crew of four, with provisions for a crew of as many as seven, and payloads. It can remain in orbit nominally for seven days and up to 30 days with special payloads. The crew controls the launch, orbital maneuvering, atmoshperic entry, and landing phases of the mission. The crew also performs payload handling.

Shuttle Performance (KSC Launch)

Figure 15 shows the maximum payload weight that can be placed into a circular orbit from Kennedy Space Center (KSC) as a function of placement orbit altitude and inclination. The payload weights shown are based on the configuration used in the reference missions and on orbital maneuvers that are limited to a Hohmann transfer from apogee of a 50- by 100-nautical mile ascent ellipse to the payload placement altitude, circularization at this altitude, and a direct descent deorbit maneuver. A 22-fps orbital reserve is carried for contingency operations.

The integral orbital maneuvering system propellant tanks aboard the orbiter carry only enough propelland to permit an empty orbiter to ascend to, and return from, an altitude of approximately 300 nautical miles. This altitude is reduced to approximately 225 nautical miles when the orbiter carries its maximum payload. During ascent, main engine cutoff occurs when the orbiter is still at a suborbital velocity. The engine cutoff location and velocity govern the impact geometry of the external tank. OMS propellant is then used to complete the ascent. Low-inclination missions launched from KSC nominally require 100 fps from the OMS to attain the required orbital velocity.

Up to three sets of auxiliary tanks can be added in the cargo bay to permit the orbiter to carry a payload to, and then descend from, higher altitude circular orbits. The additional 5,614 kilograms (12,400 pounds) of propellant in each kit, plus the 1328-kilogram (2928-pound) dry weight of the first tank kit installation, and the weights of the second and third kits, must be subtracted from the total payload capability to determine true payload weight.

Shuttle Performance VAFB Launch

Figure 16 shows that the maximum payload weight that can be delivered into a circular orbit from VAFB differs from that attainable with a KSC launch. The principal causes of this difference are the higher orbital inclinations of the missions projected from VAFB and the different range to external tank disposal. The effect of these differences in an increase in the quantity of OMS propellant expended to accelerate the orbiter to the required orbital velocity. At KSC, only 100 fps of velocity is provided by the OMS. For launches from VAFB, the OMS velocity contribution must increase to at least 350 fps just to inject the orbiter into a 50- by 100-nautical mile orbit. The increase in OMS propellant requirements is reflected as a decrease in maximum payload capability. Auxiliary OMS kits must be employed to increase the payload delivered to altitudes above 175 nautical miles.

Shuttle/IUS Performance, High Energy Orbiter

One of the potential stage vehicles to launch GPS from Shuttle is the IUS. Spacecraft and satellites whose missions require very high, super-orbital velocities (e.g., 12-hour) will employ on additional propulsive stages which must be delivered to low earth orbit, in combination with the spacecraft, by the Space Shuttle.

INERTIAL UPPER STAGE

The IUS is presently in the design definition phase. Several configurations are being considered. Of interest to the GPS application is the two-stage configuration which consists of a first stage that includes structure, electric power, and a solid rocket motor with 9,707 kilograms (21400 pounds) of propellant capacity, and a second stage which includes structure for interfacing with the payload (2.9 meters - 9.5 feet - in diameter) and with the first stage (2.3 meters - 7.6 feet - in diameter), avionics and electric power, reaction control system and propellant, and a solid rocket motor with 2,722 kilograms (6000 pounds) of propellant capacity. The overall length is 5 meters (16.4 feet). This configuration is illustrated in Figure 17, which also depicts the location of the principal components.

The design goals for the IUS are the ability to place a 2268-kilogram (5000-pound) payload into a 24-hour geosynchronous orbit, and a 2,722-kilogram (6000-pound) payload into a 12-hour, 63° orbit, when starting from the 150-nautical mile altitude of the Shuttle orbiter parking orbit.

A preliminary weight breakdown for the IUS is shown in Table 1.

Preliminary estimates of the propellant specific impulse indicate around 290 seconds.

SHUTTLE/IUS DEPLOYMENT OF GPS SATELLITE

Phase III of the GPS program will require three orbital planes with a nodal separation of 120 degrees. In each plane, eight evenly spaced GPS satellites will be deployed. The initial emplacement will require multiple satellite carried to parking orbit in each Shuttle flight. After the system is totally deployed and operational, system maintenance will require that satellites be replaced in case of malfunction. The high performance of the IUS allows two or more GPS satellites to be carried between Shuttle parking orbit to GPS operational orbit in each Shuttle flight.

A typical deployment concept that has been investigated consists of two IUS vehicles each carrying two GPS spacecraft in one Shuttle flight. This would allow deployment of two GPS satellites to each of two different orbits. The optimum Shuttle parking orbit inclination is 42.4 degrees, which results in a total Shuttle cargo weight that remains within the Shuttle capability and minimizes the time required for regression of the parking orbit from the initial nodal crossing (injection of first IUS) to the nodal crossing that allows injection of the second IUS into its operational orbit. The optimum Shuttle parking orbit inclination and schedule of IUS injection requires careful consideration of nodal regression rates, waiting time in orbit, Shuttle performance to various orbital inclinations, and visibility to remote tracking stations.

Table 1. IUS Weight Breakdown

Element	Weight, kg (lb)
Stage 1	
Burnout weight	1,177 (2,596)
Propellant	9,710 (21,408)
Expended inerts	96 (211)
Gross Weight	10,983 (24,215)
Stage 2	
Burnout weight	813 (1,793)
Propellant	2,722 (6,002)
RSC propellant	88 (193)
Expended inerts	45 (100)
Gross Weight	3,668 (8,088)
Total Vehicle Weight	14,651 (32,303)

Installation of the two IUS/GPS vehicles in the Shuttle cargo bay is illustrated in Figure 18, which also shows the weight of the vehicles, their center of gravity, and the total cargo center of gravity. Exact details of the cradle that supports the flight vehicles to the Shuttle cargo bay and the deployment procedures after arrival in orbit have been studied but not defined finally. Some of the interfaces that have been preliminarily identified are shown in Figure 19.

Figure 1. Atlas Staging Concept

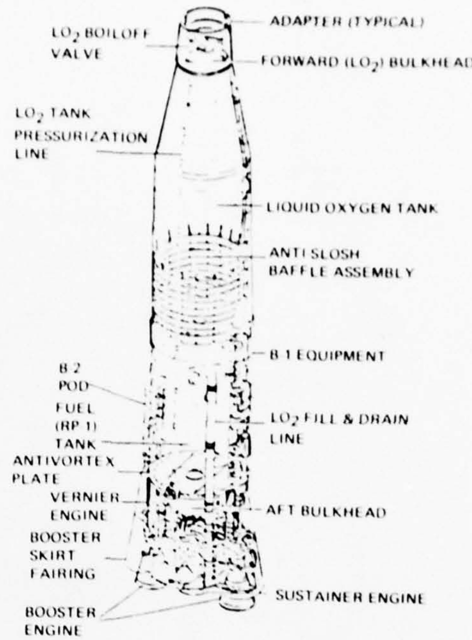


Figure 2. Atlas F Vehicle

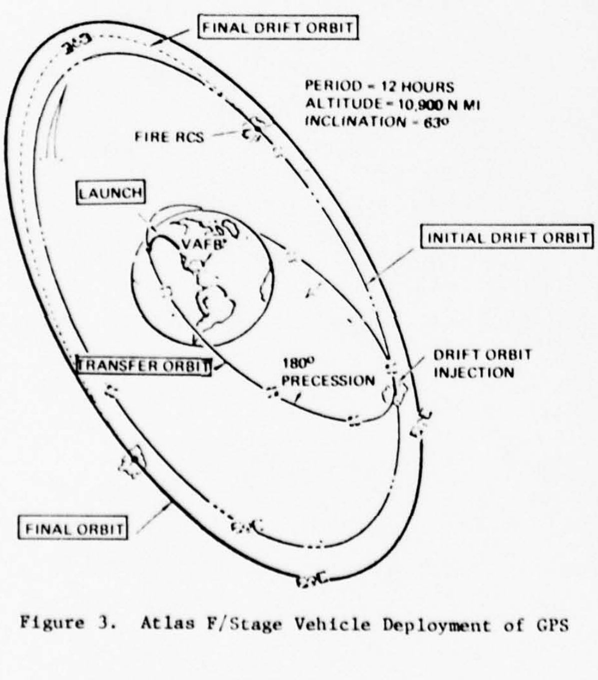


Figure 3. Atlas F/Stage Vehicle Deployment of GPS

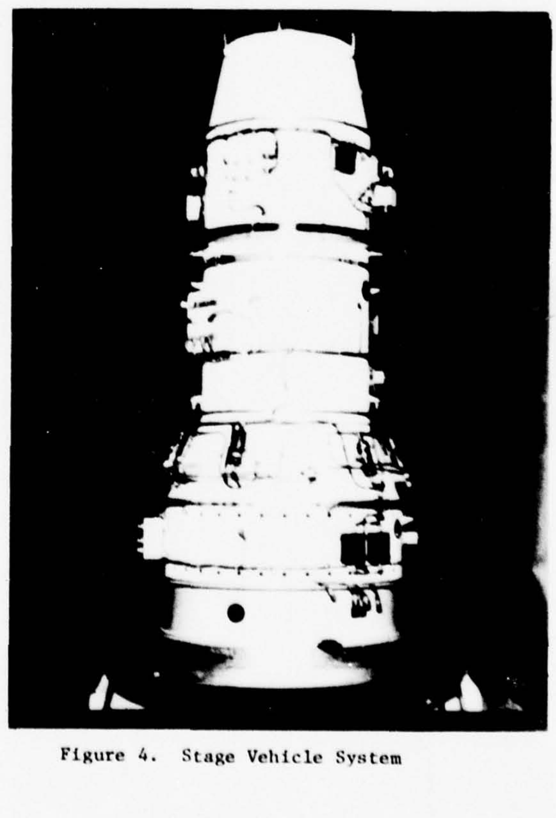


Figure 4. Stage Vehicle System

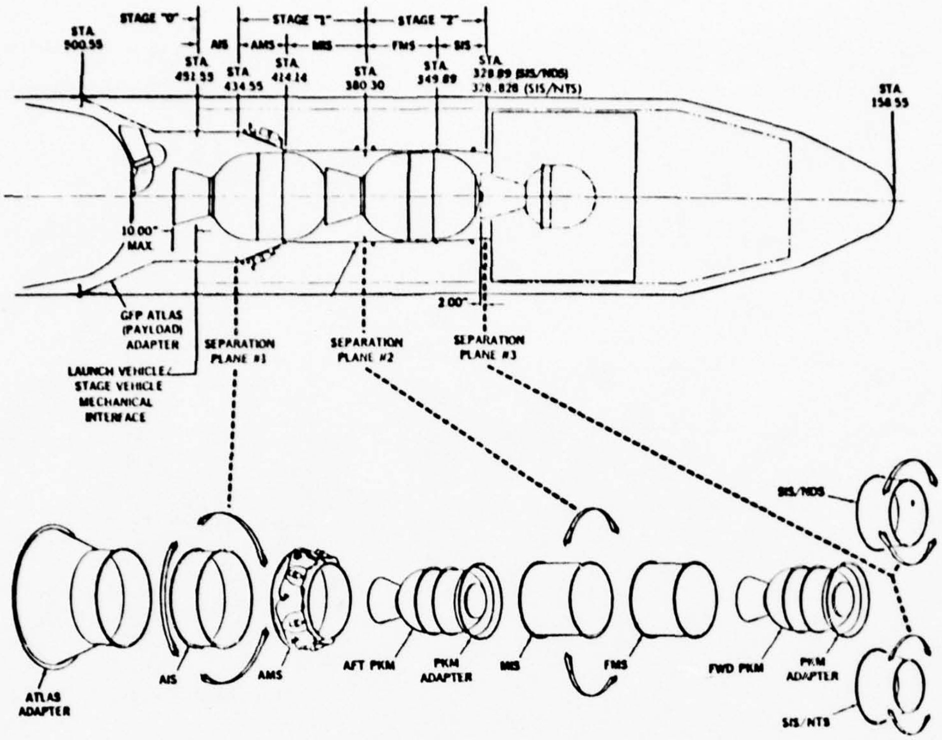


Figure 5. Stage Vehicle System

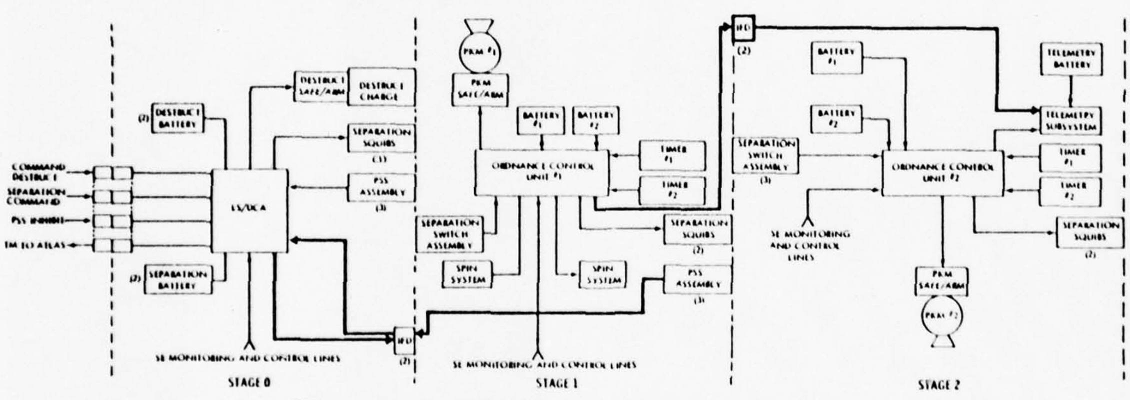


Figure 6. Stage Vehicle Functional Block Diagram

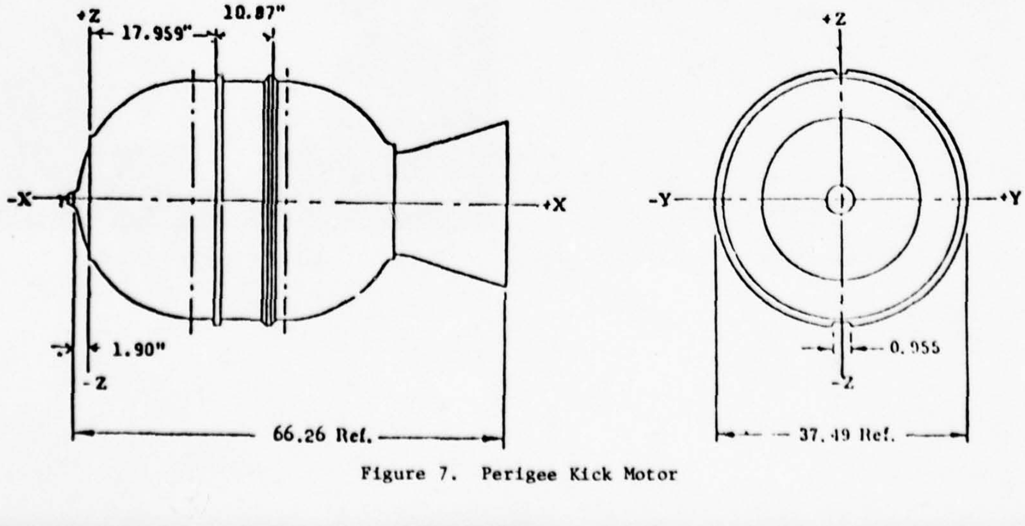


Figure 7. Perigee Kick Motor

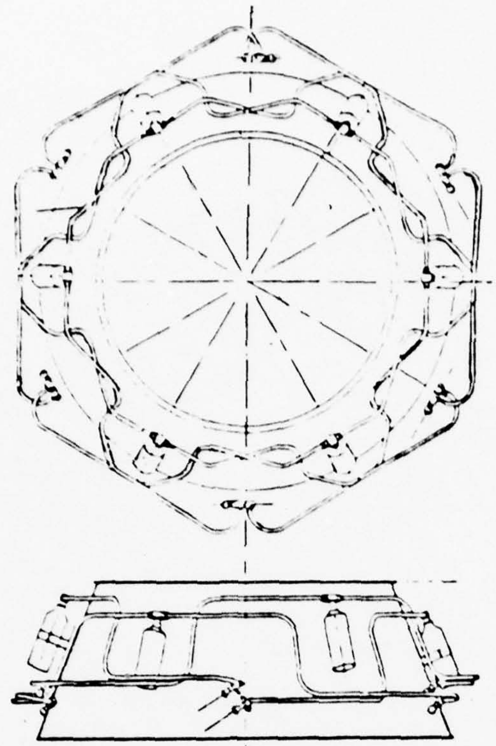


Figure 8. Stabilization System Configuration

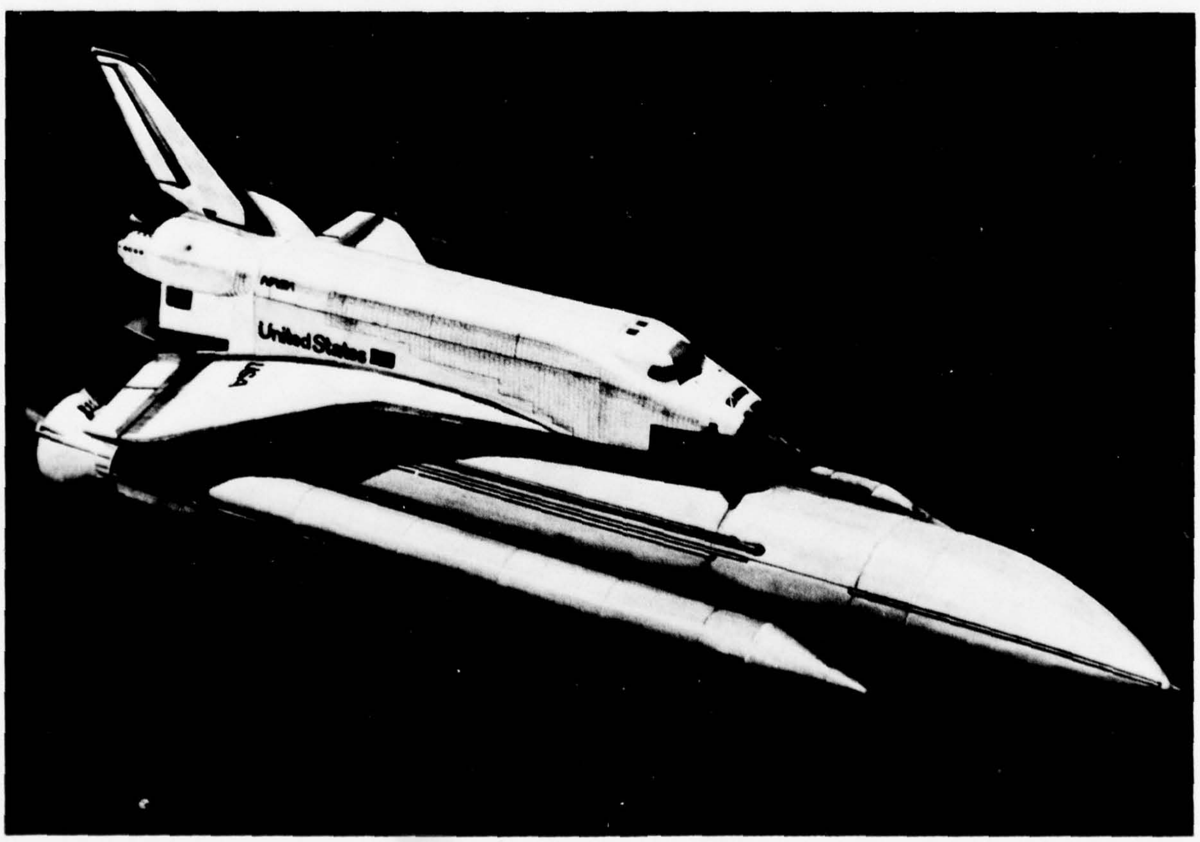


Figure 9. Space Shuttle Vehicle

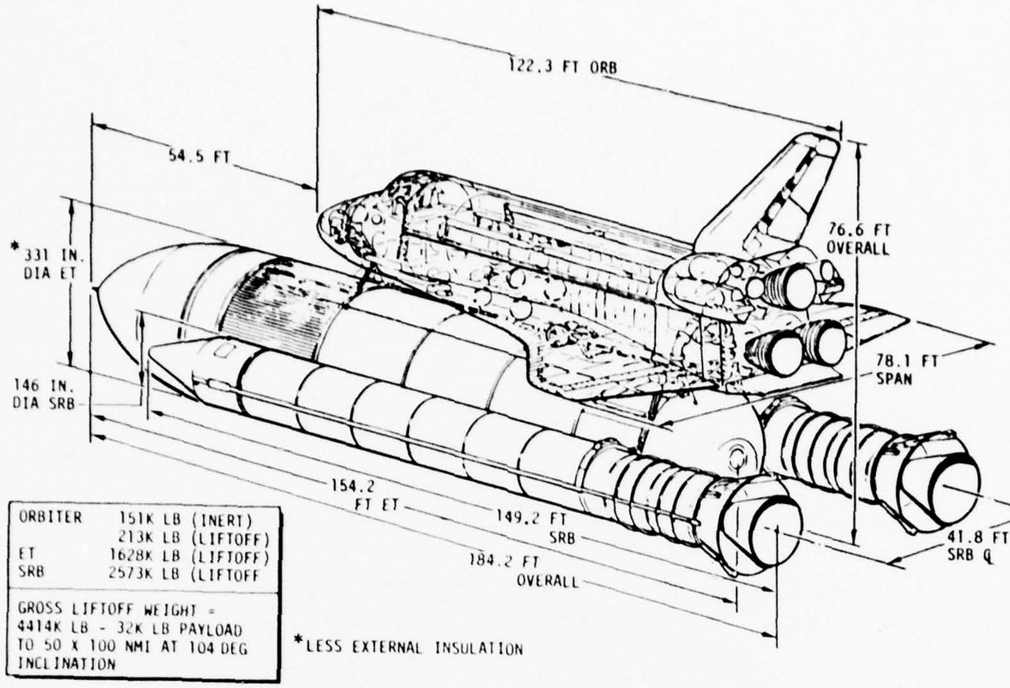


Figure 10. SSV Dimensions and Weights

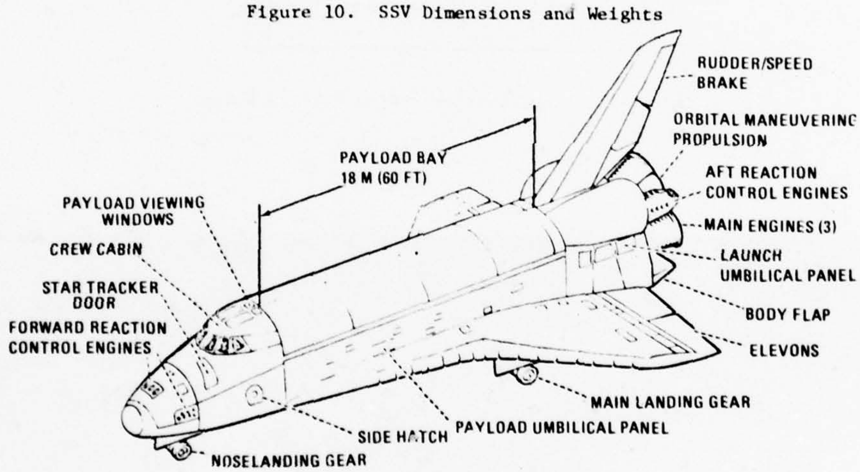


Figure 11. Space Shuttle Orbiter

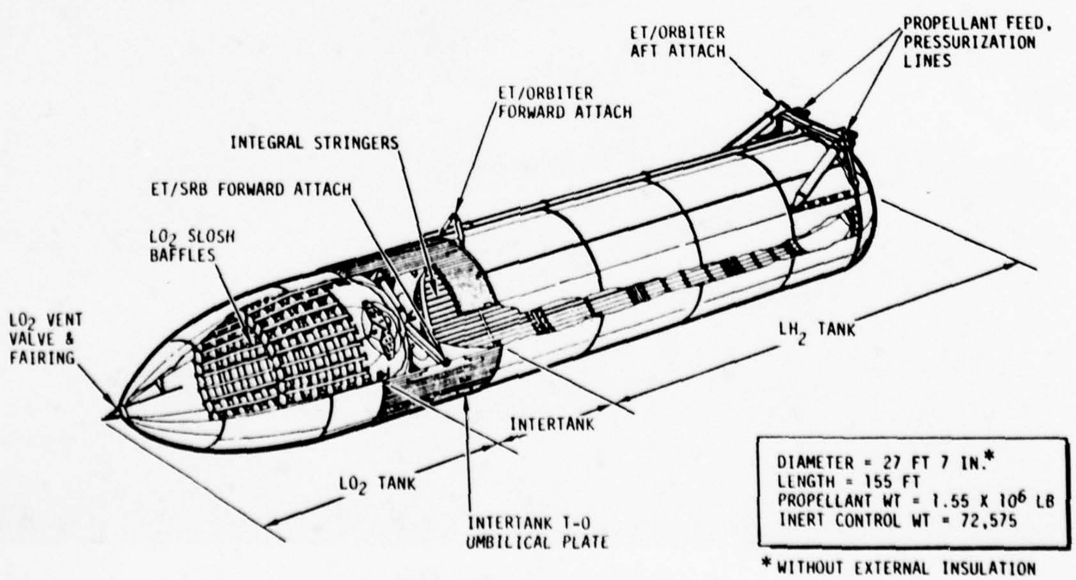


Figure 12. External Tank

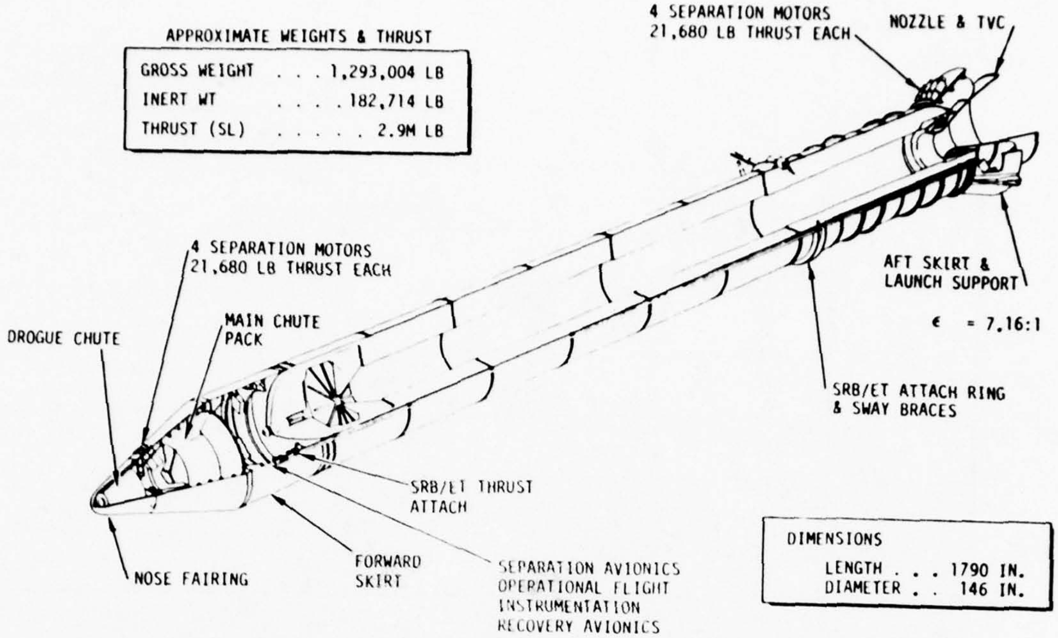


Figure 13. Solid Rocket Booster

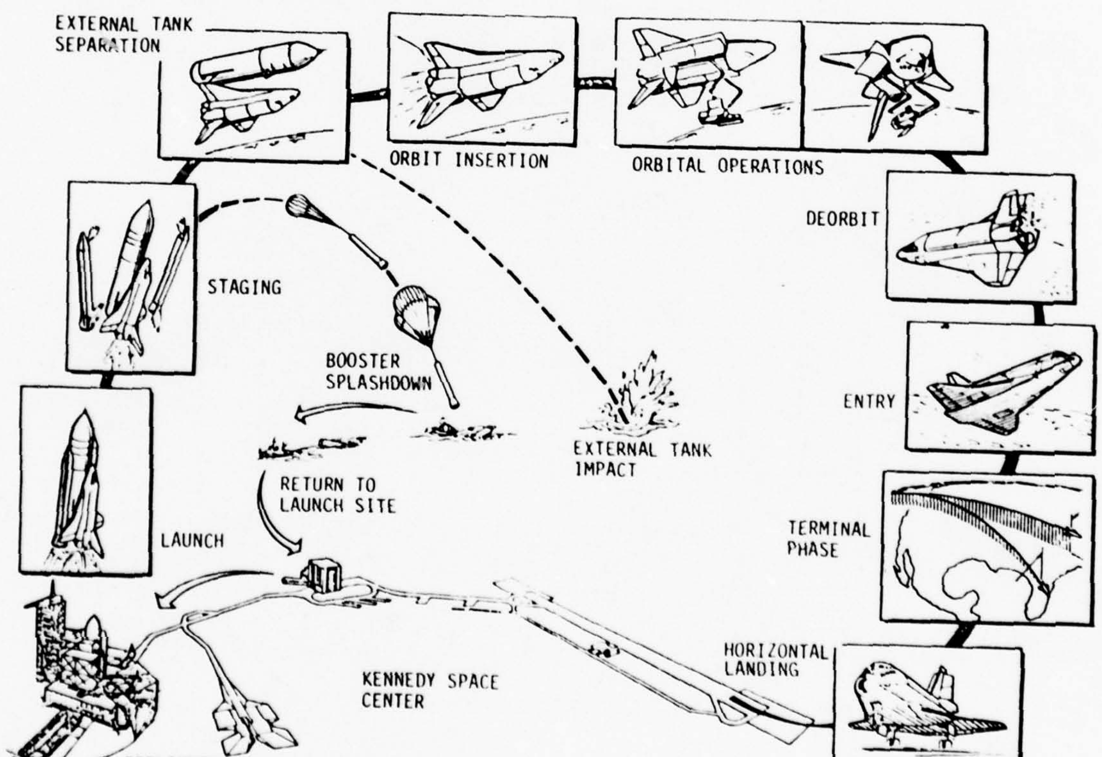


Figure 14. Mission Profile

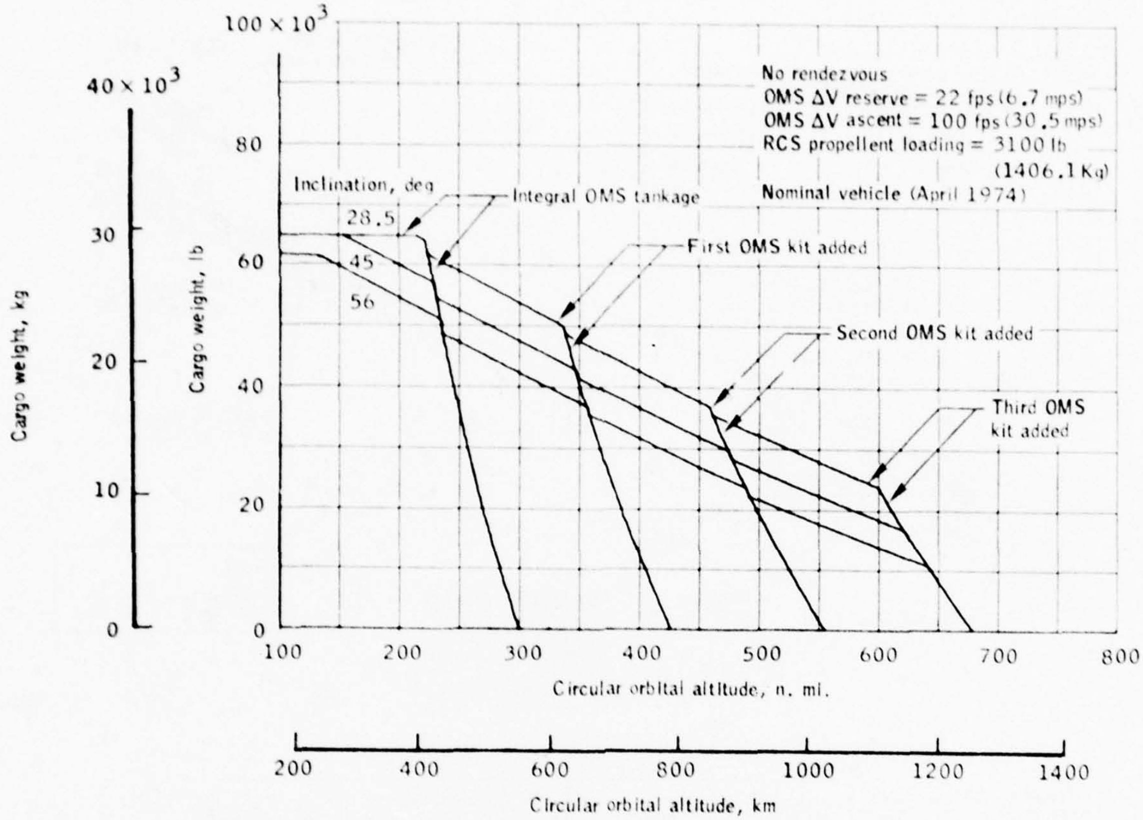


Figure 15. Payload to Circular Orbit - KSC Launch

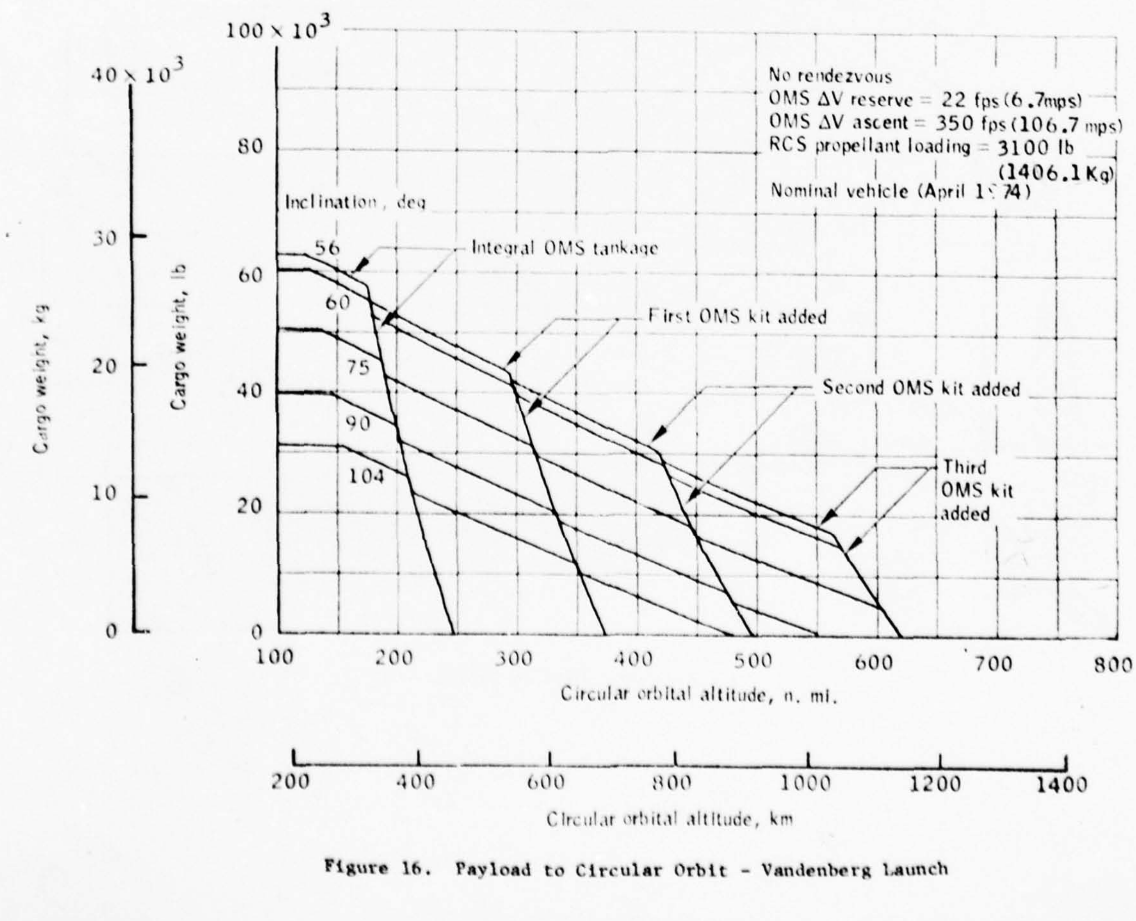


Figure 16. Payload to Circular Orbit - Vandenberg Launch

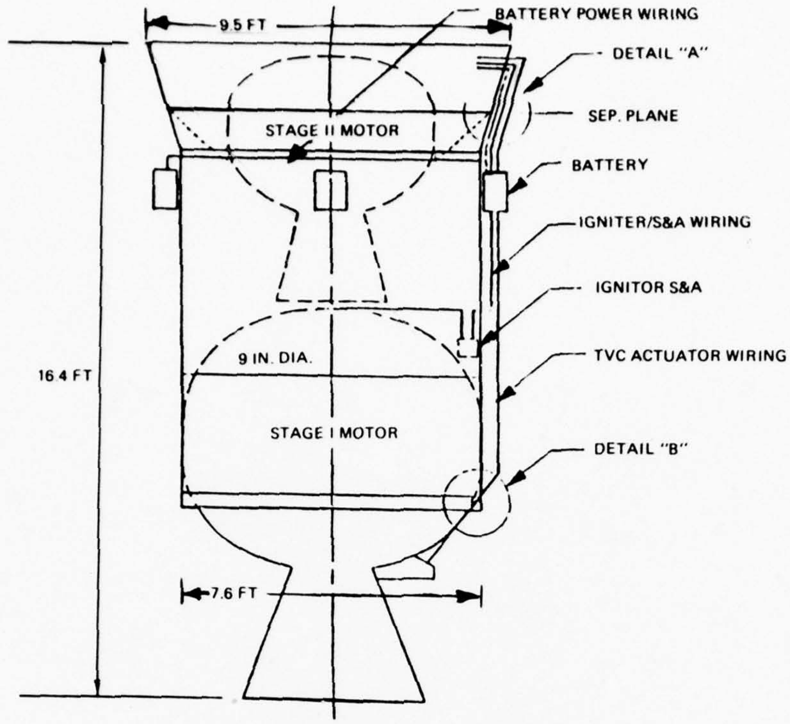


Figure 17. Two-Stage IUS

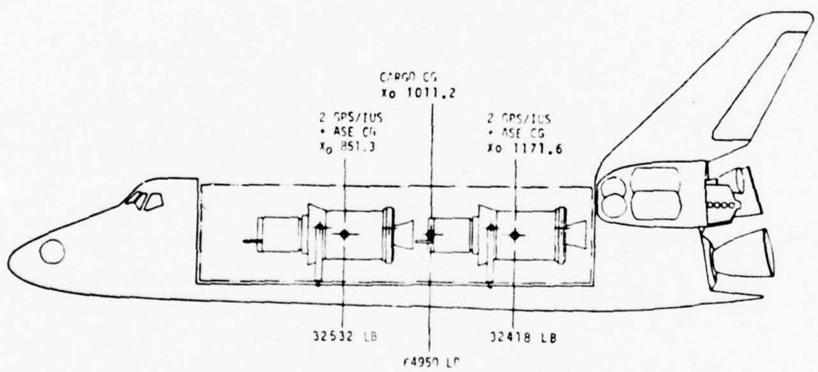


Figure 18. IUS/GPS Installation in Shuttle Cargo Bay

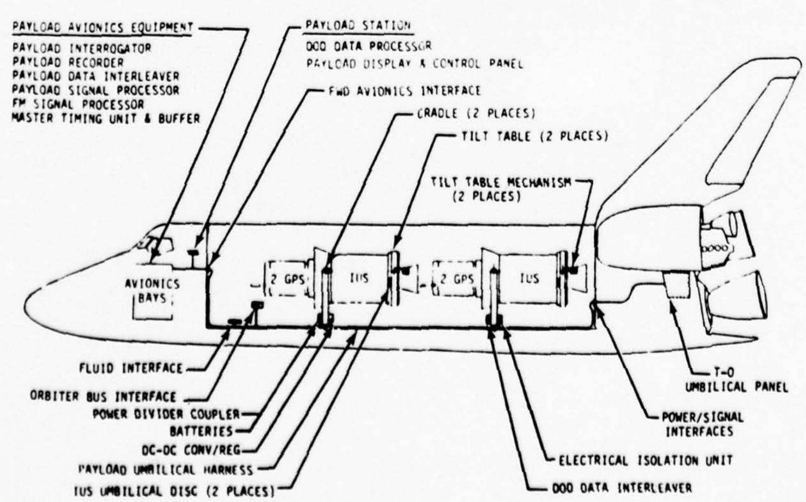


Figure 19. IUS/GPS Interfaces With Shuttle

ALTERNATE CONSTELLATIONS FOR THE GLOBAL POSITIONING SYSTEM

Hosam E. Emara-Shabaik
School of Engineering and Applied Science
7620 Boelter Hall
University of California
Los Angeles, California 90024
U.S.A.

SUMMARY

In this chapter different satellite constellations have been studied comparatively with the base line constellation considering both coverage and precision requirements. Satellite constellations have been found to be better than the base line in terms of coverage and performance of the Global Positioning System (GPS). The study is concluded with general directions and guidelines to be followed for designing constellations for the GPS.

1.0 INTRODUCTION

1.1 System Description

The Global Positioning System is a navigation system employing satellites and radio signals to provide highly accurate three dimensional fixes continuously all over the globe. The baseline constellation for this system consists of three orbital planes with eight satellites in each plane (3 x 8). Satellites in each plane are uniformally distributed in circular orbits whose inclination angles are chosen to be 63 and whose ascending nodes are uniformly separated in longitude.

The tradeoff study presented here involving the orbital parameters has shown that increasing the orbital period provides better satellite visibility and higher accuracy of navigation [1,2,4]. Also increasing the number of orbital planes is thought to be more effective in gaining better performance than increasing the number of satellites per plane [3]. Orbital inclination angle is found to effect the geographical location of better covered areas [2], specifically, higher inclination angles will move that area towards higher lattitudes. The effect of the inclination on the navigation precision in terms of PDOP has been indicated [3,4]. The navigation technique and instrumentation capabilities impose a constraint in terms of an elevation mask angle below which a satellite cannot be seen. Of course the higher the elevation mask angle the more satellites per orbit are required for global coverage [1,2], and at the same time the ability to navigate as accurately will be less for the same satellite constellation [3,4].

1.2 The Navigation Technique

Navigation in the GPS system [5,6] is based on measuring the ranges between satellites and the user. Generally speaking, measurement of range can be accomplished by timing the two-way transmission of a suitable signal originating at the user and repeated back to the user by the satellite or by comparing the time of arrival of a signal with a clock at the user synchronized to a clock at the satellite. The latter method is preferred since it eliminates radio transmission by the user and ensures a system which cannot be saturated as the number of users increases, and hence it is employed for the NAVSTAR GPS system. To maintain clock synchronization with any clock presently feasible for inclusion in the user's equipment, one additional measurement must be made. Since a three-dimensional fix is required, signals from four satellites need to be received. By measuring the time of arrival of the four signals relative to the user's clock three position coordinates and a clock correction can thus be determined. If, in addition, the range rate of change and the frequency shift are measured, the three-component velocity vector of the user can be found. A system using this measurement technique is termed a one-way pseudo-range system.

A set of inertial Earth centered coordinates, X, Y, and Z as shown in Figure (1) is introduced in which the satellite positions are assumed to be known (by the aid of the tracking stations) and in which the user wishes to obtain his position and velocity. Each satellite transmits a signal to the user. These signals contain identifiable range codes modulated upon the carrier, typically by biphase modulation. The signals also contain the equivalent of satellite ephemerides modulated at a low data rate. The signals are either at different carrier frequencies or are modulated by orthogonal codes in order that the signals may be distinguished by the user.

A similar signal is generated by the user's unsynchronized clock. By means of a correlation detector, the time shift between a satellite signal and the user clock signal is determined. This time shift (T) consists of the speed of light transit delay from the satellite to the user and the lag of the user's time reference relative to system time.

Four T_j's as measured from the four satellites together with the positions of the four satellites can be expressed as a set of four non-linear algebraic equations in four unknowns (three components of user position and the clock bias). Since the signals from the four satellites are synchronized as one of the ground station's functions, the user clock bias is the same in each of these equations. The form of these equations is as follows:

$$T_j C = \sqrt{(X-X_j)^2 + (Y-Y_j)^2 + (Z-Z_j)^2} - \Delta \cdot C \qquad (1)$$

where

C = speed of light
T_j = Time shift of the signal from satellite No. j

X_j, Y_j, Z_j = Components of the position vector of the jth satellite

X, Y, Z = Components of the position vector of the user

Δ = The lag of the user clock

$j = 1, 2, 3, 4$

If the location of the user relative to the satellites allows the determination of the four T_j's and the satellite geometry does not produce algebraic singularities, the equations can then be solved[1] for the user position and clock bias.

If, at the same time as the pseudo ranges are measured, the rate of change is measured by normal Doppler extraction techniques, four equations with user velocity and user clock frequency bias as unknowns may be written and solved for user velocity in three dimensions and user clock frequency bias. The equations for velocity determination are as follows:

$$C\dot{T}_j = \frac{(X-X_j)(\dot{X}-\dot{X}_j) + (Y-Y_j)(\dot{Y}-\dot{Y}_j) + (Z-Z_j)(\dot{Z}-\dot{Z}_j)}{\sqrt{(X-X_j)^2 + (Y-Y_j)^2 + (Z-Z_j)^2}} - C\dot{\Delta} \qquad (2)$$

where

$\dot{X}, \dot{Y}, \dot{Z}$ = The user velocity components

$\dot{\Delta}$ = User clock frequency bias

1.2. Circular Orbits

In this chapter satellite constellations with circular orbits are considered. A circular orbit is defined by its radius "R_o" which is directly related to the orbital period by the following relationship.

$$T = \frac{2\pi (R_o)^{\frac{3}{2}}}{(G\ m_e)^{\frac{1}{2}}} \qquad (3)$$

where

T = The orbital period in seconds.

G = Universal gravitational constant

$= 6.673 \times 10^{-8}$ dyne-cm^2/gm^2

m_e = Earth's mass

$= 5.977 \times 10^{27}$ gm-mass

Denoting the earth's mean radius by "R_e", which is the radius of a sphere of volume equal to the earth's volume, then

$$R_e = 6371.2 \text{ Km} = 6371.2 \times 10^5 \text{ cm}$$

Equation (3) can be rewritten in terms of the orbit to earth radii ratio (R_o/R_e) as

$$T = \frac{2\pi (\frac{R_o}{R_e})^{\frac{3}{2}} (R_e)^{\frac{3}{2}}}{(G\ m_e)^{\frac{1}{2}}} \qquad (4)$$

$$(\frac{R_o}{R_e})^{\frac{3}{2}} = \frac{(G\ m_e)^{\frac{1}{2}}\ T}{2\pi (R_e)^{\frac{3}{2}}}$$

$$= \frac{3600\ (G\ m_e)^{\frac{1}{2}}\ T}{2\pi (R_e)^{\frac{3}{2}}} \qquad (5)$$

where T in this last expression is to be substituted in hours. Substituting the values for different terms in (5) we get

$$(\frac{R_o}{R_e}) = 0.797\ T^{\frac{2}{3}} \qquad (6)$$

where

T = the orbital period in hours.

A circular orbit is fixed in the inertial frame by defining the longitude of its ascending node (Ω), (which is the longitude of the point a satellite in orbit crosses the equatorial plane going from the southern to the northern part of the orbit) and its inclination angle (i), (which is the angle measured upward from the equatorial plane to the orbital plane).

2.0 MATHEMATICAL DESCRIPTION OF THE PROBLEM

In this formulation of the problem, all positions are taken in reference to an inertial earth centered frame as shown in Fig. (1).

2.1. User's Position

A user's position on the earth can be defined as shown in Fig. (2) by two angles ($\theta_{\ell g}$), and ($\theta_{\ell t}$) where

$\theta_{\ell g}(t)$ = The angle measured to the user's longitude at time (t)

$\theta_{\ell t}(t)$ = The angle measured to the user's latitude at time (t).

Then,

$$\theta_{\ell g}(t) = \theta_{\ell g}(t_o) + \omega_e (t-t_o) \tag{7}$$

$$\theta_{\ell t}(t) = \theta_{\ell t}(t_o) \tag{8}$$

where

ω_e = the earth's angular velocity
$= \frac{2\pi}{24}$ rad./hr.

t = time in hours.

t_o = the initial time when observation starts.

The position vector to a user $\bar{R}_u$ can be expressed as

$$\bar{R}_u = \begin{bmatrix} X_u \\ Y_u \\ Z_u \end{bmatrix} = R_e \begin{bmatrix} \cos\theta_{\ell t} \cos\theta_{\ell g} \\ \cos\theta_{\ell t} \sin\theta_{\ell g} \\ \sin\theta_{\ell t} \end{bmatrix} \tag{9}$$

Figure (3) illustrates this expression.

2.2 Satellite's Position

As mentioned earlier the orbits considered are circular with radius (R_o) and their ascending nodes are uniformly spaced in longitude. Also the satellites are uniformly spaced in the orbit. Let $\bar{R}_{mn}$ be the position vector to the nth satellite in the mth orbital plane. Then as shown in Figure (4) we can write.

$$\bar{R}_{mn} = \begin{bmatrix} X_{mn} \\ Y_{mn} \\ Z_{mn} \end{bmatrix} = R_o \begin{bmatrix} \sin\phi_{mn}(t) \cos i \cos\Omega_m - \cos\phi_{mn}(t) \sin\Omega_m \\ \sin\phi_{mn}(t) \cos i \sin\Omega_m + \cos\phi_{mn}(t) \cos\Omega_m \\ \sin\phi_{mn}(t) \sin i \end{bmatrix} \tag{10}$$

$$\Omega_m = \frac{m-1}{M} 2\pi \tag{11}$$

$$\phi_{mn}(t) = \phi_{mn}(t_o) + \frac{2\pi}{T}(t-t_o) \tag{12}$$

$$\phi_{mn}(t_o) = \frac{n-1}{N} 2\pi + \frac{m-1}{M}\frac{2\pi}{N} \tag{13}$$

where

Ω_m = the longitude of the ascending node of the mth orbit.

M = number of orbital planes

m = order of an orbital plane
= 1, 2, ..., M

i = the orbital inclination angle.

$\phi_{mn}(t_o)$ = the initial angle from the ascending node to the satellite in the direction of motion.

N = number of satellites per plane

n = order of a satellite in a plane
= 1, 2, ..., N

2.3 Satellite Visibility

Considering the user-satellite geometry shown in Figure (5).

$$\bar{U} = \bar{R}_u - \bar{R}_{mn} \tag{14}$$

$$\bar{U}\cdot\bar{R}_u = -|\bar{R}_u||\bar{U}| \cos \beta \tag{15}$$

$$\cos \beta = - \frac{\bar{U}\cdot\bar{R}_u}{|\bar{R}_u||\bar{U}|} = - \hat{R}_u \cdot \hat{U} \tag{16}$$

$$\beta = 90^o - \alpha \tag{17}$$

$$\hat{R}_u = \frac{\bar{R}_u}{|\bar{R}_u|} \tag{18}$$

= unit vector in the direction of $\bar{R}_u$, i.e., components of $\hat{R}_u$ are the direction cosines of the user.

$$\hat{U} = \frac{\bar{U}}{|\bar{U}|} \tag{19}$$

= unit vector in the direction from user to satellite.

α = user - satellite elevation angle.

If the elevation mask angle is taken to be (α_m), then the nth satellite in the mth orbit will be visible by the user if

$$0^o \leq \beta \leq 90^o - \alpha_m \tag{20}$$

2.4 Geometric Performance

The geometric performance of the GPS is measured in terms of the position dilution of precision (PDOP) [15]. PDOP is the ratio of the position error standard deviation to that of the pseudo range error under the assumption that the pseudo range errors from different satellites are uncorrelated with equal standard deviations. Smaller value of PDOP indicates better geometric performance in the sense that the observer's position will be more accurately determined.

Refer to the nth satellite in the mth orbit as the ith satellite for simplicity of notations in what follows then,

$$\bar{D}_i = \bar{R}_i - \bar{R}_u \tag{21}$$

Let

$$\bar{e}_i = \bar{D}_i/D_i \tag{22}$$

and

$$\bar{e}_u = \bar{R}_u/R_u \tag{23}$$

Then

$$D_i\bar{e}_i = \bar{R}_i - \bar{R}_u \tag{24}$$

Taking the dot product of equation (24) with $\bar{e}_i$ then

$$D_i = \bar{R}_i\cdot\bar{e}_i - \bar{R}_u\cdot\bar{e}_i \tag{25}$$

Assuming that the observer tracks the range signal perfectly and that there are no system bias, then the apparent range ρ_i, called the pseudo range to the ith satellite is given by

$$\rho_i = D_i + B_i + B_u \tag{26}$$

where

B_u = Bias caused by the difference between the user clock and the system time.

B_i = Bias caused by the difference between the satellite clock and the system time.

Then from (25), (26)

$$\rho_i = \bar{R}_i\cdot\bar{e}_i - \bar{R}_u\cdot\bar{e}_i + B_i + B_u \tag{27}$$

$$\bar{R}_u\cdot\bar{e}_i - B_u = \bar{R}_i\cdot\bar{e}_i - \rho_i + B_i \tag{28}$$

The above equaiton can be used for computing the user position and time bias ($\bar{R}_u$ and B_u). Considering the range measurements to four different satellites, four equations of the form (28) can be written in a matrix form as follows.

$$\begin{bmatrix} \bar{e}_1^T & \vline & -1 \\ \bar{e}_2^T & \vline & -1 \\ \bar{e}_3^T & \vline & -1 \\ \bar{e}_4^T & \vline & -1 \end{bmatrix} \begin{bmatrix} \bar{R}_u \\ \hline B_u \end{bmatrix} = \begin{bmatrix} \bar{e}_1^T \; 1 & & & \\ & \bar{e}_2^T \; 1 & & \\ & & \bar{e}_3^T \; 1 & \\ & & & \bar{e}_4^T \; 1 \end{bmatrix} \begin{bmatrix} \bar{R}_1 \\ B_1 \\ \bar{R}_2 \\ B_2 \\ \bar{R}_3 \\ B_3 \\ \bar{R}_4 \\ B_4 \end{bmatrix} - \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} \tag{29}$$

We can write the above matrix equation in a compact form as

$$G\bar{X} = A\bar{S} - \rho = \bar{Y} \tag{30}$$

where

$$G = \begin{bmatrix} \bar{e}_1^T & \vline & -1 \\ \bar{e}_2^T & \vline & -1 \\ \bar{e}_2^T & \vline & -1 \\ \bar{e}_3^T & \vline & -1 \end{bmatrix} , \quad \bar{X} = \begin{bmatrix} \bar{R}_u \\ \hline B_u \end{bmatrix}$$

$$A = \begin{bmatrix} \bar{e}_1^T \; 1 & & & \\ & \bar{e}_2^T \; 1 & & \\ & & \bar{e}_3^T \; 1 & \\ & & & \bar{e}_4^T \; 1 \end{bmatrix}$$

$$\bar{S} = \begin{bmatrix} \bar{R}_1 \\ B_1 \\ \bar{R}_2 \\ B_2 \\ \bar{R}_3 \\ B_3 \\ \bar{R}_4 \\ B_4 \end{bmatrix} , \quad \bar{\rho} = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix}$$

Equation (30) is a system of linear equations with four unknown (the elements of $\bar{X}$). The measurement vector $\bar{Y}$ generally involves measurement errors.

Considering the measurement error vector $\bar{V}$, then equation (30) becomes

$$\bar{Y} = G\bar{X} + \bar{V} \tag{31}$$

The estimate $\hat{\bar{X}}$ for $\bar{X}$ is given by [12].

$$\hat{\bar{X}} = [G[\,cov^{-1}(\bar{V})]G]^{-1}\; G^T[cov^{-1}(\bar{V})]\bar{Y} \tag{32}$$

The estimation error is given as $\delta\bar{X} = \bar{X} - \hat{\bar{X}}$ (33)

and the error covariance matrix is given as

$$\mathrm{cov}\ (\delta\bar{X}) = E\{\delta\bar{X}\ \delta\bar{X}^T\} = [G^T[\mathrm{cov}^{-1}(V)]G]^{-1} \qquad (34)$$

Equation (34) gives the covariance of the errors in position and time that an observer would have if he precisely knew the measurement error covariance [cov(V)]. It can be said that [cov($\bar{V}$)] reflects the precision of the GPS instruments, while G and G^T reflect the geometrical precision of performance of the GPS. Regardless of some geometric effects included in [cov(V)] such as the effect of the elevation angle on ionosphere modeling errors, as a good approximation the geometric performance can be evaluated by considering [cov(V)] equal to the identity matrix in equation (34). Consequently the geometric dilution of precision (GDOP) is defined as

$$\mathrm{cov}(\delta\bar{X}) = (G^T\ G)^{-1} \qquad (35)$$

That is if

$$(G^T\ G)^{-1} = \begin{bmatrix} \delta_{11} & \delta_{12} & \delta_{13} & \delta_{14} \\ \delta_{21} & \delta_{22} & \delta_{23} & \delta_{24} \\ \delta_{31} & \delta_{32} & \delta_{33} & \delta_{34} \\ \delta_{41} & \delta_{42} & \delta_{43} & \delta_{44} \end{bmatrix} \qquad (36)$$

Then, the position dilution of precision is given by

$$\mathrm{PDOP} = \sqrt{\delta_{11} + \delta_{22} + \delta_{33}} \qquad (37)$$

Similarly the time dilution of precision is given by

$$\mathrm{TDOP} = \sqrt{\delta_{44}} \qquad (38)$$

2.5 Satellites Selection

This section discusses the selection of four satellites among those visible in the sky to provide the lowest PDOP and hence the highest navigation accuracy.

Referring to Figure 6, the tetrahedron formed by the unit vectors from the user to the four satellites is defined by

$$\bar{A} = \bar{e}_2 - \bar{e}_1 \qquad (39)$$

$$\bar{B} = \bar{e}_3 - \bar{e}_1 \qquad (40)$$

$$\bar{C} = \bar{e}_4 - \bar{e}_1 \qquad (41)$$

and its volume is expressed as

$$V = \frac{1}{6}\ \bar{C} \cdot (\bar{A}\mathrm{x}\bar{B}) \qquad (42)$$

It can be shown that there is a high correlation between the volume of the tetrahedron and PDOP [15]. Namely PDOP is inversely proportional to the volume of the tetrahedron. Thus, the best four satellites to select from among those available in the sky are those which possess the biggest tetrahedron volume. Moreover it can be proven that the highest satellite in the sky is one of those best four satellites. Hence, the final decision to make is to select the best three satellites available which when combined with the highest one in the sky will form the beggest tetrahedron. The highest satellite in the sky is the one that has the smallest "β" angle among all visible satellites.

3.0 ALTERNATE CONSTELLATIONS FOR COVERAGE

The requirement that a GPS user be able to fix his position in three dimensions with the possibility of clock bias, necessitates the simultaneous observation of at least four satellites at any moment in time from any user location on the globe. The baseline GPS orbit configuration consists of 24 satellites deployed in three circular, 63 inclined, subsynchronous, 12-hour orbits. Eight satellites are uniformly distributed in each of three orbit planes whose nodes are separated by 120° in longitude.

The results presented here are those arrived at by Emara in [1], and [2] aiming at finding satellite constellations that provide three dimensional continuous global coverage with a number of satellites less than 24 (the base line design number). In [1] the preliminary studies were done considering the ratio of the orbital radius to the earth radius ($\frac{R_o}{R_e}$) and the orbital inclination angle (i) as the decision variables. The elevation mask angle, (α_m) was considered as a constraint imposed by the

navigation technique and instrumentations available. The outcome of the study was expressed as a recommended constellation with (M) orbital planes and (N) satellites per plane, (MxN).

In [2], a more extensive study was done aiming at analyzing the geographical characteristics of coverage in terms of the geographical distribution of satellites availability. Again as in [1], $(\frac{R_o}{R_e})$, (i) were considered the decision variables, (α_m) was taken as a constraint. The outcome was presented as recommended satellite constellations with their associated geographical availability distribution. The geographical availability distribution was tabulated as time averages of the number of satellites observed at different locations. A comparison was made between different constellations from which general conclusions were drawn as guidelines to decide on a specific constellation.

3.1 Results

The optimal satellite constellations[1] specified in terms of the number of planes "M" times the number of satellites per plane "N" and the orbital inclination angle are given in Tables 1 and 2 for different combinations of orbital periods and mask angles. For the case of three orbital planes a satellite constellation of five satellites per plane at an inclination angle of 45.0°, 63.0°, or 75.0° and an orbital period of 48.0 or 72.0 hours satisfies the coverage requirements[2] with a minimum number of satellites for both 5.0° and 10.0° mask angles.[3] The effect of changing the inclination angle on the satellite observability can be noticed from the average observability statistics given in Tables 3, 4, and 5 for the 48.0 hour period and from Tables 6, 7, and 8 for the 72.0 hour period. These tables provide the time average of the number of satellites observable at different locations. A constellation of 4 x 5 with a period of 24 hours and an inclination of 45.0° or with a period of 48 or 72 hours and an inclination of 45.0°, 63.0°, or 75.0° satisfies the coverage requirements in both cases of 5.0° and 10.0° mask angles as shown by Table 2. The average observability statistics of these constellations are given in Tables 9 through 15.

3.2 Conclusions

From the results of this section we can make several conclusions:

1. Comparing the results in Table 1 with those in Table 2 we find that the three plane configuration is preferred to the four plane configuration. It allows a smaller number of satellites to be used.

2. Table 1 shows that the opitmal constellation that satisfies the coverage requirements under the limitation of both 5.0° and 10.0° elevation mask angles consists of three planes with five satellites per plane and a period of either 48.0 or 72.0 hours and an inclination angle of 45.0°, 63.0°, or 75.0°.

3. From Tables 3 through 8 and Tables 10 through 15 by comparing the average observability statistics of the same satellite constellation at different inclination angles we conclude that increasing the orbit inclination angle shifts the region where more satellites are available (on the average) away from the equator.

4. From Tables 3 through 8 and Tables 10 through 15 by comparing the average observability statistics of two constellations under similar conditions one of three orbital planes and the other of four orbital planes we see that the four configuration provides more visible satellites on the average at different locations. This may be due to the fact that more satellites are being used in the four orbital planes than in the three orbital planes.

5. Examining Tables 3 through 15 again, but this time we compare the average observability statistics of the same constellation operating under different orbital periods, we conclude that when a constellation is satisfying the coverage requirements with a certain period will give more visible satellites on the average when used with a longer orbital period. Another way to put this conclusion is to say that the number of satellites required for three-dimensional, continuous, worldwide coverage decreases as the orbit radius is increased.

6. There appears to be no general trend regarding the effect of the inclination angle on the number of satellites required for coverage.

7. We should keep in mind that in this work optimal constellations are found based on the condition of satisfying the coverage requirements. Other considerations such as precision of performance and redundancy in case of satellite failure would require a higher number of satellites than that found here.

4.0 ALTERNATE CONSTELLATIONS FOR PERFORMANCE

One of the objectives of the GPS is to provide extremely accurate three-dimensional position fixes. The degree of navigation accuracy by means of satellites is substantially coupled with the choice of orbit configuration and is directly related to the choice of satellites among those visible in the sky.

---

1. A satellite constellation that requires the least number of satellites to provide continuous three dimensional global coverage is called an optimal stacllite constellation.

2. The coverage requirements are mentioned in the introduction of this part.

3. Note that according to the definition of the mask angle, a satellite constellation satisfying the coverage requirements for 10 mask angle will consequently satisfy the coverage requirements for 5° mask.

The measure of the geometric performance of the GPS is the Position Dilution of Precision (PDOP) which was defined earlier in Sec. (2.4) eq. (37). PDOP was found to be inversely proportional to the volume of the tetrahedron formed by the four satellites chosen to compute the position fix and the time bias [15]. Hence, a user should choose the four satellites among those visible in the sky which will provide the biggest volume of the tetrahedron. One of those satellites is the highest one in the sky (the one that has the smallest angle "β").

In [3], [4], Yonezawa studies the geometric performance of different satellite constellations as compared to that of the baseline configuration of 3 x 8. *His results are those presented here.* The study suggests some new constellations which provide better geometric performance than the baseline constellation.

4.1 Results

In [3] performance characteristics were studied for the following constellations: (3 x 8), (4 x 6), (5 x 4), (5 x 5), (6 x 3), (6 x 4), (7 x 3), (7 x 4), (8 x 3). The results are presented in terms of an average value of PDOP in Figures (7) through (24). The latitude vs. performance curves for each constellation are drawn at the inclination angle which acquires that constellation its best performance inclination curves.

The baseline 3 x 8 configuration has a significantly large PDOP average value about 45 degrees in latitude for the elevation mask of 10 degrees, as shown in Figure (8). This divergence of PDOP average value is caused by the fact that at times only four satellites are visible to the observer about latitude 45 degrees in the case of 10 degree mask. On the other hand, there are no such divergences for 10 degree mask in 4 x 6 (Figure 10), 5 x 5 (Figure 14), 6 x 4 (Figure 18), 7 x 3 (Figure 20), 7 x 4 (Figure 22), and 8 x 3 (Figure 24) due to satellite visibility of at least five as shown in Table (16).

If global latitude performances in 5 degree mask for different constellations are compared with each other, it may be said that (5 x 5), (6 x 4), (8 x 3) are better than (3 x 8); (4 x 6), (7 x 4) are as good as (3 x 8); (5 x 4), (6 x 3), (7 x 3) are worse than (3 x 8). Based on the performance in both cases of 5° and 10° mask, (6 x 4) and (7 x 4) are the best and (4 x 6), (7 x 3), and (5 x 5) come in at second place. Noting that (7 x 4) required more satellites than the baseline we are left with (6 x 4) as the best choice.

In [4] the performance of the following constellations, (3 x 8) baseline, (4 x 6), (6 x 4) is studied at different orbital periods and mask angles. The geometric performance is considerably improved as the orbit period increases. For the elevation mask ranging form 5° to 10° the geometric performance at a specified inclination is improved by making orbit period longer. However, no further substantial improvement may be possible if orbit period reaches to a certain value, say 24 hours, (Figures 27, 31, 35). For higher elevation mask enhancement in geometric performance and satellite visibility can be achieved by increasing orbit period to some extent (Figures 27, 31, 35). The number of visible satellites for each configuration at several elevation mask angles is shown statistically in Tables (17, 18, 19). It may be said that four visible satellites chosen through the satellite selection algorithm stay longer as an optimal set for longer orbit period.

4.2 Conclusions

1. Comparing the performance of different constellations in the case of 5° mask we can say that (5 x 5), (6 x 4) and (8 x 3) are better than the baseline while (4 x 6) and (7 x 4) are as good as the baseline and (5 x 4) and (7 x 3) are worse than (3 x 8), and finally (6 x 3) is not worth consideration.
2. Looking at the performance curves for 10° mask we can rank the different constellations as follows: (6 x 4), (7 x 4) are the best and (4 x 6), (7 x 3), (5 x 5) come in at second place. Noting that (7 x 4) requires more satellites than the baseline, then the best choice is (6 x 4).
3. Increasing the number of orbits may be more effective in enhancing the geometric performance of the orbit configuration than increasing the number of satellites in each orbit. (For example, (6 x 4) is better than (4 x 6) and (8 x 3) is better than (3 x 8)).[1]
4. The following constellations, (3 x 8) at 63° inclination, (4 x 6) at 50° inclination, and (6 x 4) at 60° inclination provide very much the same precision when employed with a 24 hour orbital period.

REFERENCES

1. Emara, Hosam E., and C. T. Leondes, "Minimum Number of Satellites for Three-Dimensional Continuous Worldwide Coverage," IEEE Trans. Aerospace and Electronic Systems, Vol AES-13, March 1977.

2. Emara, Hosam E., and C. T. Leondes, "Coverage Analysis for the Global Positioning System," a UCLA report, under preparation.

---

1. To allow for multiple launches using the shuttle, having more than three orbits is not preferred, a constraint which was not stated at the time of developing these results.

3. Yonezawa, K., "Geometric Performance of Global Positioning System," UCLA-ENG-7576, October 1975.

4. Yonezawa, K., "Geometric Performance of Global Positioning System with Respect to Orbit Period," UCLA-ENG-76 , May 1976.

5. Keane, L.M., "A Multiple User Satellite System for Navigation and Traffic Control," NASA, Electronic Research Center, Cambridge, Mass. 02139, EASCON 1969 Record, pp. 190-197.

6. Woodford, J. B., W. C. Melton, and R. L. Dutcher, "Satellite Systems for Navigation Using 24-hour Orbits," The Aerospace Corporation, EASCON 1969 Record, pp. 184-189.

7. Haviland, Robert P., and C. M. House, Handbook of Satellites and Space Vehicles, D. Van Nostrand Company, Inc.

8. Ehricke, K. A., Space Flight, Vol. II: Dynamics, D. Van Nostrand Company, Inc.

9. Stuhlinger, E. and G. Mesmer, Space Science and Engineering, McGraw-Hill Book Company.

10. Thomson, W. T., Introduction to Space Dynamics, John Wiley, New York, 1961.

11. Berman, A. I., The Physical Principles of Astronautics, John Wiley and Sons, New York, 1961.

12. Leibelt, P. B., An Introduction to Optimal Estimation, Ch. 5, Addison-Wesley, Reading, Mass., 1967.

13. Soule, P. W., "Preliminary Mathematical Considerations for System 621B Position Determination Equations and Error Analysis," TOR-1001 (2527-17)-6, Aerospace Corporation, El Segundo, CA, August 1967.

14. Soule, P. W., "A Static Error Analysis of a Class of Navigation Systems Employing Satellites," TOR-0158 (3525)-12, Aerospace Corporation, El Segundo, CA, January 1968.

15. Bogen, A. H., "Geometric Performance of the Global Positioning System," TR-0074 (4461-02)-1, Aerospace Corporation, El Segundo, CA, June 1974.

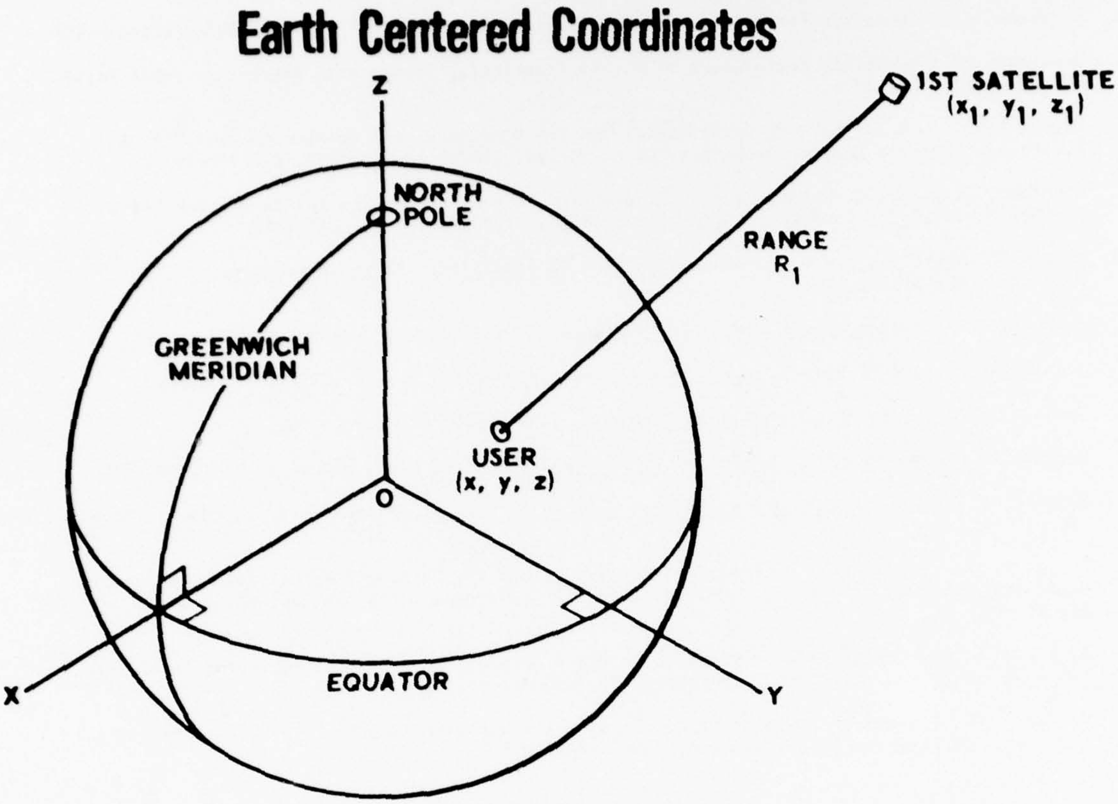


Fig.1 Earth centered coordinates

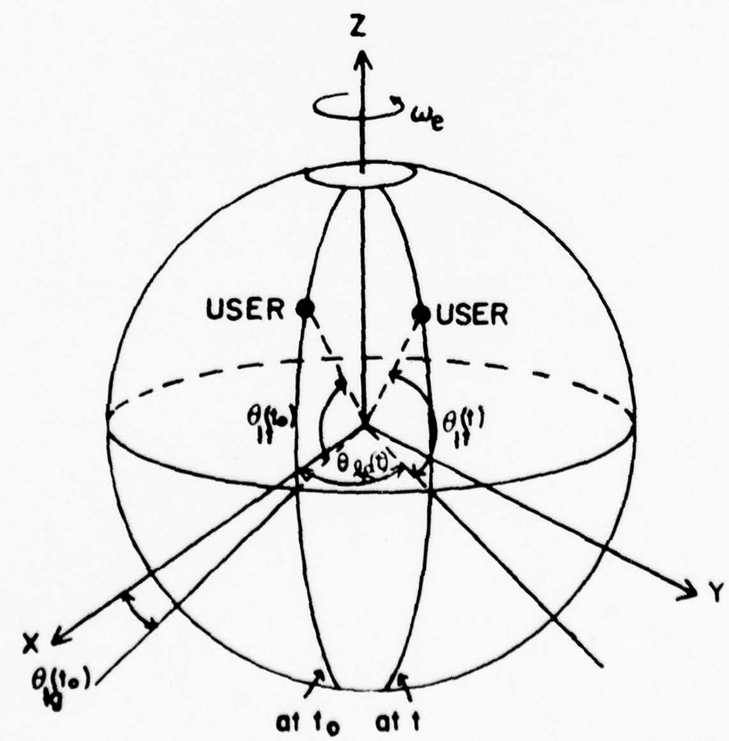


Fig.2 Geometrical configuration for User position in terms of longitude and latitude

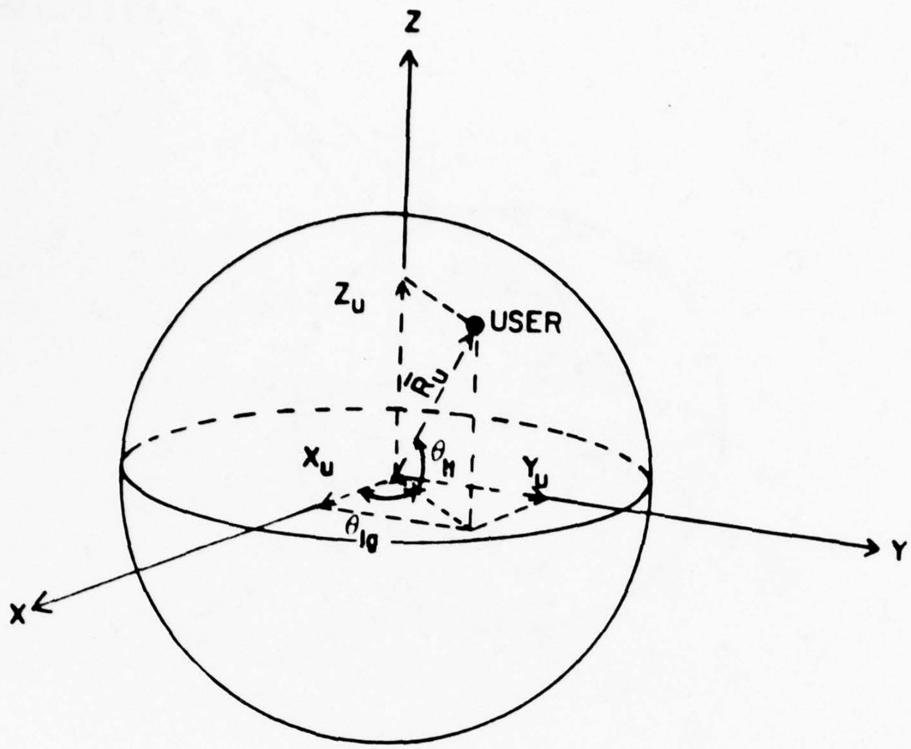


Fig.3 Geometrical configuration for User Position in Cartesian coordinates

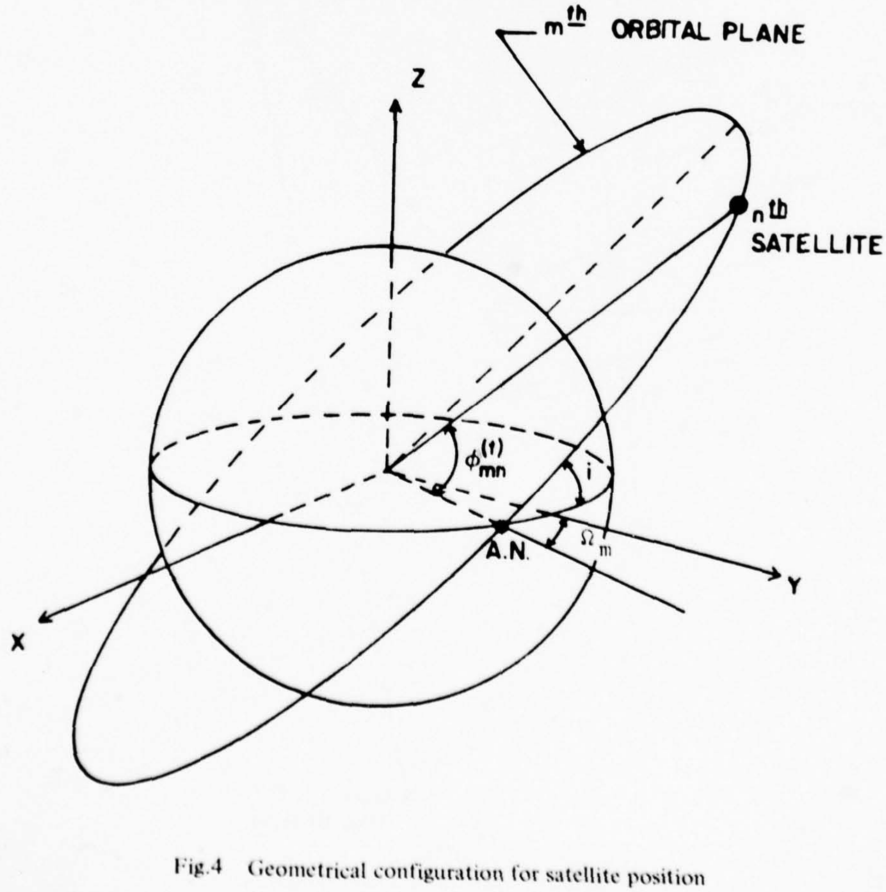


Fig.4 Geometrical configuration for satellite position

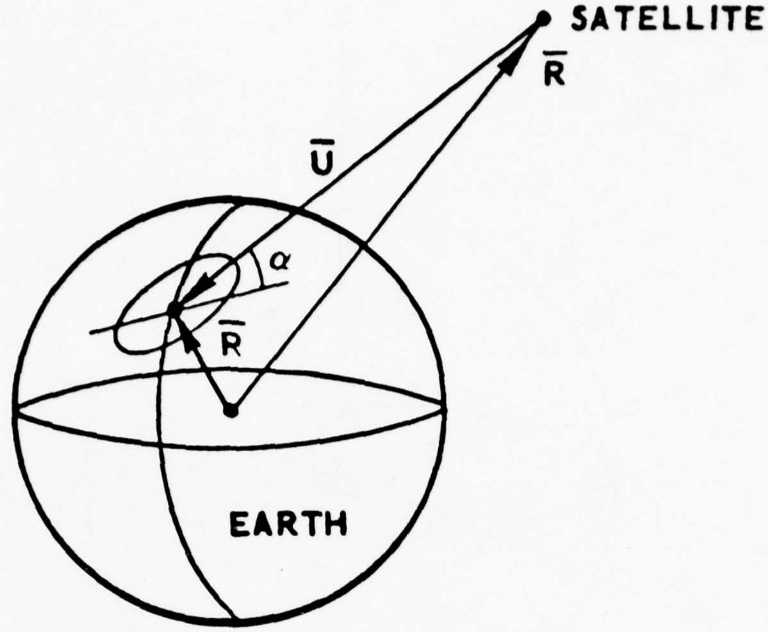


Fig.5 User-satellite geometry

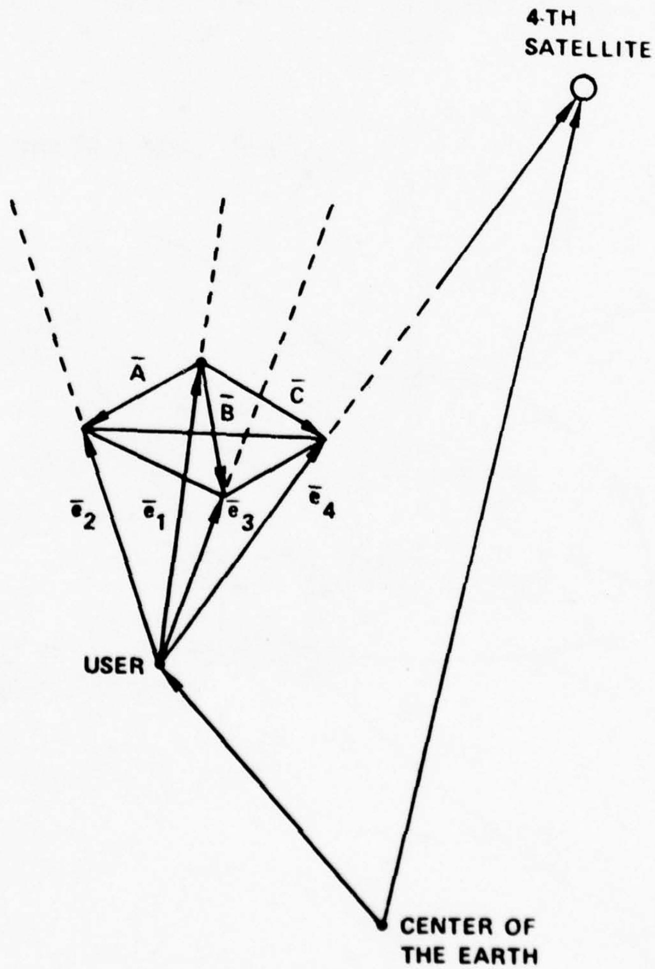


Fig.6 Tetrahedron formed by 4 satellites

Table 1

Satellite Constellations for Three Orbital Planes

Mask Angle	Constellation	Period in Hours			
		12.0	24.0	48.0	72.0
5.0°	M×N	3×5	3×5	3×5	3×5
	Inclination Angle	63.0°	45.0° or 63.0° or 75.0°	30.0° or 45.0° or 63.0° or 75.0°	30.0° or 45.0° or 63.0° or 75.0°
10.0°	M×N	3×6	3×6	3×5	3×5
	Inclination Angle	45.0° or 63.0° or 75.0°	30.0° or 45.0° or 63.0° or 75.0°	45.0° or 63.0° or 75.0°	45.0° or 63.0° or 75.0°

Table 2

Satellite Constellations for Four Orbital Planes

Mask Angle	Constellation	Period in Hours			
		12.0	24.0	48.0	72.0
5.0°	M×N	4×5	4×4	4×4	4×4
	Inclination Angle	63.0°	75.0°	45.0° or 75.0°	45.0° or 75.0°
10.0°	M×N	4×6	4×5	4×5	4×5
	Inclination Angle	63.0° or 75.0°	45.0°	45.0° or 63.0° or 75.0°	45.0° or 63.0° or 75.0°

Table 3

Average Observability Statistics for the 3×5 Constellation
with 48 hr-period, 10° Mask, 45° Inclination

Latitude \ Longitude	0°	45°	90°	135°	180°	225°	270°	315°
90°	5.69							
60°	5.12	5.25	4.87	5.25	5.12	4.81	5.69	4.81
30°	5.31	5.37	5.62	5.37	5.31	5.5	5.31	5.5
0°	6.06	6.00	5.94	6.00	6.06	6.00	5.94	6.00
-30°	5.31	5.5	5.31	5.5	5.31	5.37	5.62	5.37
-60°	5.12	4.81	5.69	4.81	5.12	5.25	4.87	5.26
-90°	5.69							

Table 4

Average Observability Statistics for the 3×5 Constellation
with 48 hr-period, 10° Mask, 63° Inclination

Latitude \ Longitude	0°	45°	90°	135°	180°	225°	270°	315°
90°	6.06							
60°	5.81	5.87	5.81	5.87	5.81	5.81	6.19	5.81
30°	5.06	5.06	5.62	5.06	5.06	5.12	5.06	5.12
0°	5.94	5.56	5.56	5.56	5.94	5.56	5.56	5.56
-30°	5.06	5.12	5.06	5.12	5.06	5.06	5.62	5.06
-60°	5.81	5.81	6.19	5.81	5.81	5.87	5.81	5.87
-90°	6.06							

Table 5

Average Obersvability Statistics for the 3×5 Constellation
with 48 hr-period, 10° Mask, 75° Inclination

Latitude \ Longitude	0°	45°	90°	135°	180°	225°	270°	315°
90°	6.19							
60°	6.06	6.06	5.81	6.06	6.06	6.12	6.31	6.12
30°	5.37	5.5	5.5	5.5	5.37	5.5	5.44	5.5
0°	5.06	4.94	4.87	4.94	5.06	4.94	4.87	4.94
-30°	5.37	5.5	5.44	5.5	5.37	5.5	5.5	5.5
-60°	6.06	6.12	6.31	6.12	6.06	6.06	5.81	6.06
-90°	6.19							

Table 6

Average Observability Statistics for the 3×5 Constellation
with 72 hr-period, 10° Mask, 45° Inclination

Latitude \ Longitude	0°	45°	90°	135°	180°	225°	270°	315°
90°	5.79							
60°	5.40	5.43	5.40	5.39	5.37	5.39	5.40	5.43
30°	5.58	5.67	5.65	5.64	5.67	5.64	5.65	5.67
0°	6.08	6.11	6.08	6.11	6.08	6.11	6.08	6.11
-30°	5.58	5.67	5.65	5.64	5.67	5.64	5.65	5.67
-60°	5.40	5.43	5.40	5.39	5.37	5.39	5.40	5.43
-90°	5.79							

Table 7

Average Observability Statistics for the 3×5 Constellation
with 72 hr-period, 10° Mask, 63° Inclination

Latitude \ Longitude	0°	45°	90°	135°	180°	225°	270°	315°
90°	6.21							
60°	6.04	6.06	6.04	6.01	6.04	6.01	6.04	6.06
30°	5.25	5.26	5.25	5.29	5.26	5.29	5.25	5.26
0°	5.75	5.78	5.75	5.83	5.81	5.83	5.75	5.78
-30°	5.25	5.26	5.25	5.29	5.26	5.29	5.25	5.26
-60°	6.04	6.06	6.04	6.01	6.04	6.01	6.04	6.06
-90°	6.21							

Table 8

Average Observability Statistics for the 3×5 Constellation
with 72 hr-period, 10° Mask, 75°-Inclination

Latitude \ Longitude	0°	45°	90°	135°	180°	225°	270°	315°
90°	6.29							
60°	6.21	6.19	6.18	6.15	6.15	6.15	6.18	6.19
30°	5.65	5.65	5.65	5.62	5.60	5.62	5.65	5.65
0°	5.25	5.28	5.31	5.25	5.19	5.25	5.31	5.28
-30°	5.65	5.65	5.65	5.62	5.60	5.62	5.65	5.65
-60°	6.21	6.19	6.18	6.15	6.15	6.15	6.18	6.19
-90°	6.29							

Table 9

Average Observability Statistics for the 4×5 Constellation
with 24 hr-period, 10° Mask, 45° Inclination

Latitude \ Longitude	0°	45°	90°	135°	180°	225°	270°	315°
90°	7.17							
60°	6.25	6.12	6.25	6.25	6.12	6.25	6.12	6.12
30°	6.58	6.71	6.5	6.5	6.71	6.58	6.79	6.79
0°	7.67	7.33	7.67	7.67	7.33	7.67	7.33	7.33
-30°	6.58	6.71	6.5	6.5	6.71	6.58	6.79	6.79
-60°	6.25	6.12	6.25	6.25	6.12	6.25	6.12	6.12
-90°	7.17							

Table 10

Average Observability Statistics for the 4×5 Constellation
with 48 hr-period, 10° Mask, 45° Inclination

Latitude \ Longitude	0°	45°	90°	135°	180°	225°	270°	315°
90°	7.58							
60°	7.10	6.54	6.96	6.96	6.54	7.10	6.69	6.69
30°	7.12	7.31	7.17	7.17	7.31	7.12	7.27	7.27
0°	8.00	8.00	8.00	8.00	8.00	8.00	8.00	8.00
-30°	7.31	7.12	7.27	7.27	7.12	7.31	7.17	7.17
-60°	6.54	7.10	6.69	6.69	7.10	6.54	6.96	6.69
-90°	7.58							

Table 11

Average Observability Statistics for the 4×5 Constellation
with 48 hr-period, 10° Mask, 63° Inclination

Longitude / Latitude	0°	45°	90°	135°	180°	225°	270°	315°
90°	8.08							
60°	7.92	7.75	7.9	7.9	7.75	7.92	7.77	7.77
30°	6.75	6.98	6.77	6.77	6.98	6.75	6.96	6.96
0°	7.54	7.54	7.54	7.54	7.54	7.54	7.54	7.54
-30°	6.98	6.75	6.96	6.96	6.75	6.98	6.77	6.77
-60°	7.75	7.92	7.77	7.77	7.92	7.75	7.9	6.90
-90°	8.08							

Table 12

Average Observability Statistics for the 4×5 Constellation
with 48 hr-period, 10° Mask, 75° Inclination

Longitude / Latitude	0°	45°	90°	135°	180°	225°	270°	315°
90°	8.25							
60°	8.17	8.04	8.19	8.19	8.04	8.17	8.02	8.02
30°	7.27	7.29	7.27	7.27	7.29	7.27	7.29	7.29
0°	6.6	6.60	6.6	6.6	6.6	6.6	6.6	6.6
-30°	7.29	7.27	7.29	7.29	7.27	7.29	7.27	7.27
-60°	8.04	8.17	8.02	8.02	8.17	8.04	8.19	8.19
-90°	8.25							

Table 13

Average Observability Statistics for the 4×5 Constellation
with 72 hr-period, 10° Mask, 45° Inclination

Latitude \ Longitude	0°	45°	90°	135°	180°	225°	270°	315°
90°	7.72							
60°	7.21	7.21	7.19	7.19	7.18	7.18	7.19	7.19
30°	7.53	7.53	7.49	7.50	7.50	7.50	7.50	7.49
0°	8.17	8.17	8.17	8.17	8.17	8.17	8.17	8.17
-30°	7.53	7.53	7.49	7.50	7.50	7.50	7.50	7.49
-60°	7.21	7.21	7.19	7.19	7.18	7.18	7.19	7.19
-90°	7.72							

Table 14

Average Observability Statistics for the 4×5 Constellation
with 72 hr-period, 10° Mask, 63° Inclination

Latitude \ Longitude	0°	45°	90°	135°	180°	225°	270°	315°
90°	8.28							
60°	8.04	8.04	8.18	7.96	8.10	8.10	7.96	8.18
30°	7.00	7.00	7.03	7.01	7.04	7.04	7.01	7.03
0°	7.75	7.75	7.61	7.83	7.69	7.69	7.83	7.61
-30°	7.00	7.00	7.03	7.01	7.04	7.04	7.01	7.03
-60°	8.04	8.04	8.18	7.96	8.10	8.10	7.96	8.18
-90°	8.28							

Table 15

Average Observability Statistics for the 4×5 Constellation
with 72 hr-period, 10° Mask, 75° Inclination

Latitude \ Longitude	0°	45°	90°	135°	180°	225°	270°	315°
90°	8.39							
60°	8.24	8.24	8.35	8.18	8.29	8.29	8.18	8.35
30°	7.50	7.50	7.58	7.47	7.56	7.56	7.47	7.58
0°	7.08	7.08	6.89	7.17	6.97	6.97	7.17	6.89
-30°	7.50	7.50	7.58	7.47	7.56	7.56	7.47	7.58
-60°	8.24	8.24	8.35	8.18	8.29	8.29	8.18	8.35
-90°	8.39							

Table 16

Number of Visible Satellites (All Area)

ORBIT CONSTELLATION	INCLINATION (deg)	ELEVATION MASK 5° MIN	AVERAGE	MAX	ELEVATION MASK 10° MIN	AVERAGE	MAX
3 × 8	63	6	8.1	11	4	7.2	9
4 × 6	50	6	8.2	10	5	7.2	9
5 × 4	65	5	6.8	9	4	6.0	9
5 × 5	70	6	8.4	10	5	7.5	10
6 × 3	55	4	6.2	8	4	5.4	8
6 × 4	60	7	8.3	11	5	7.2	11
7 × 3	65	5	7.2	11	5	6.4	8
7 × 4	55	7	9.5	13	6	8.4	12
8 × 3	60	6	8.7	11	6	7.2	10

Note: Average is taken over uniformly distributed observers on the earth.

TABLE 17

Number of Visible Satellites for

63 Degree Inclined 3 × 8 Configuration

Elevation Mask	ORBIT PERIOD 12 hours Min	Average	Max	24 hours Min	Average	Max	96 hours Min	Average	Max
5°	6	8.1	11	6	9.2	11	9	10.4	12
10°	4	7.2	9	6	8.2	11	6	9.3	12
15°	4	6.3	9	4	7.3	9	6	7.2	11
20°	4	5.4	8	4	6.4	9	4	6.3	9
25°	3		7	4	5.4	8	4	5.4	8
30°	2		7	3		7	4		8
35°	1		5	2		7	3		6

NOTE: **Average is taken over uniformly distributed observers on the earth.**

TABLE 18

Number of Visible Satellites for

50 Degree Inclined 4 × 6 Configuration

Elevation Mask	ORBIT PERIOD								
	12 hours			24 hours			96 hours		
	Min	Average	Max	Min	Average	Max	Min	Average	Max
5°	6	8.2	10	6	9.2	11	6	10.2	12
10°	5	7.2	9	6	8.2	11	6	9.3	12
15°	5	6.3	8	6	7.3	9	6	8.3	11
20°	3		7	5	6.4	8	5	7.2	10
25°	2		7	3		7	5	6.3	8
30°	2		6	2		7	3		8
35°	1		6	2		6	2		7

NOTE: Average is taken over uniformly distributed observers on the earth.

TABLE 19

Number of Visible Satellites for

60 Degree Inclined 6 × 4 Configuration

Elevation Mask	ORBIT PERIOD								
	12 hours			24 hours			96 hours		
	Min	Average	Max	Min	Average	Max	Min	Average	Max
5°	7	8.3	11	7	9.2	12	9	10.4	12
10°	5	7.2	11	6	8.2	11	7	9.2	11
15°	4	6.2	9	5	7.3	11	6	8.3	11
20°	4	5.5	7	4	6.3	10	5	7.3	11
25°	3		7	4	5.5	7	4	6.4	10
30°	2		7	2		7	4	5.5	7
35°	1		6	2		7	2		7

NOTE: Average is taken over uniformly distributed observers on the earth.

Figure 7 not available at time of publication

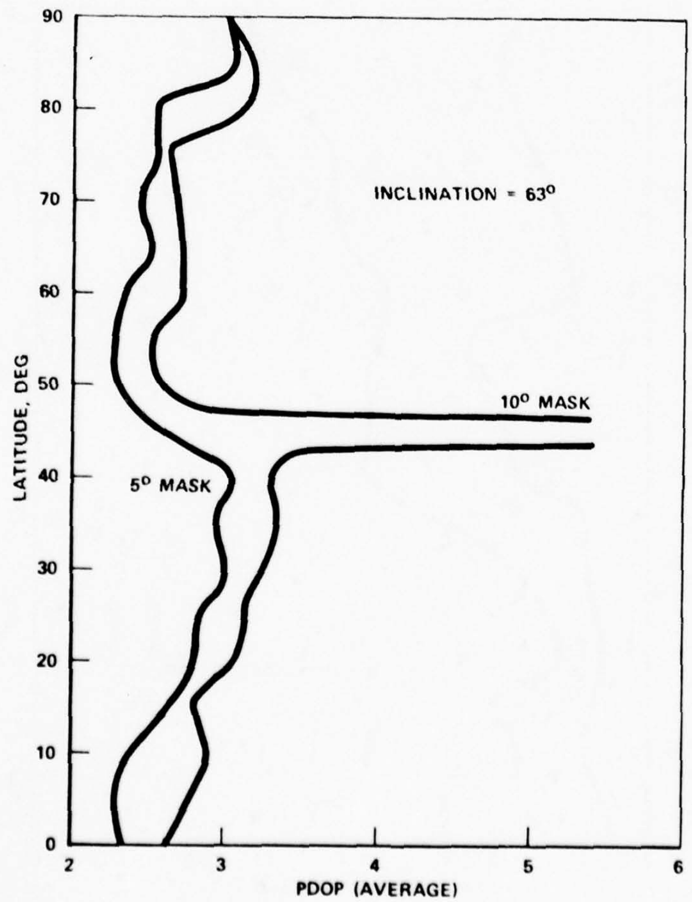


Fig.8 3 x 8 Global latitude performance

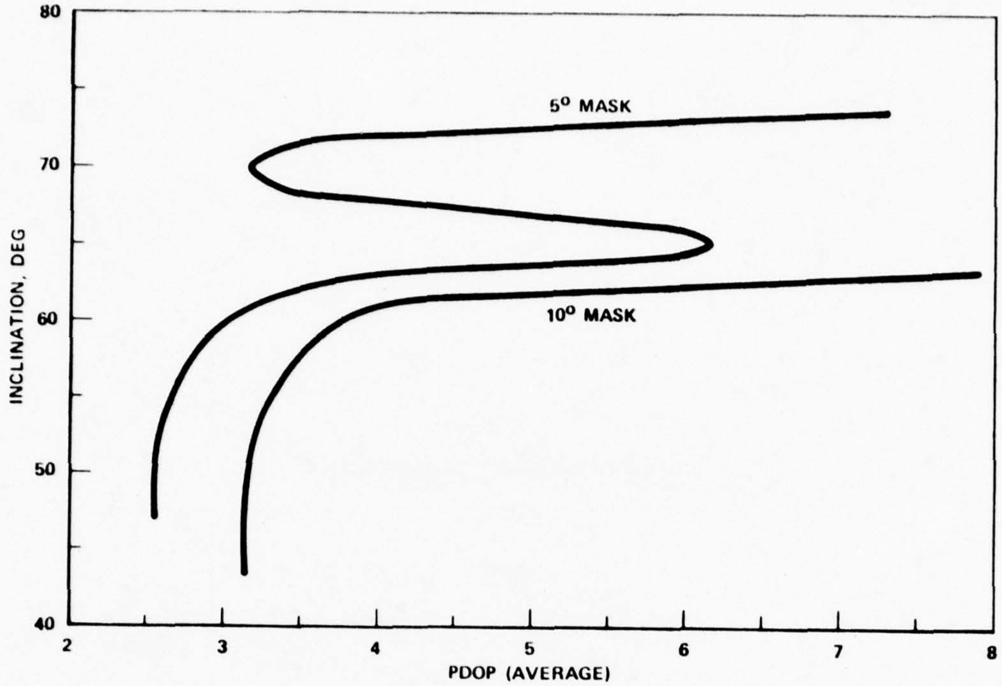


Fig.9 4 x 6 Global geometric performance

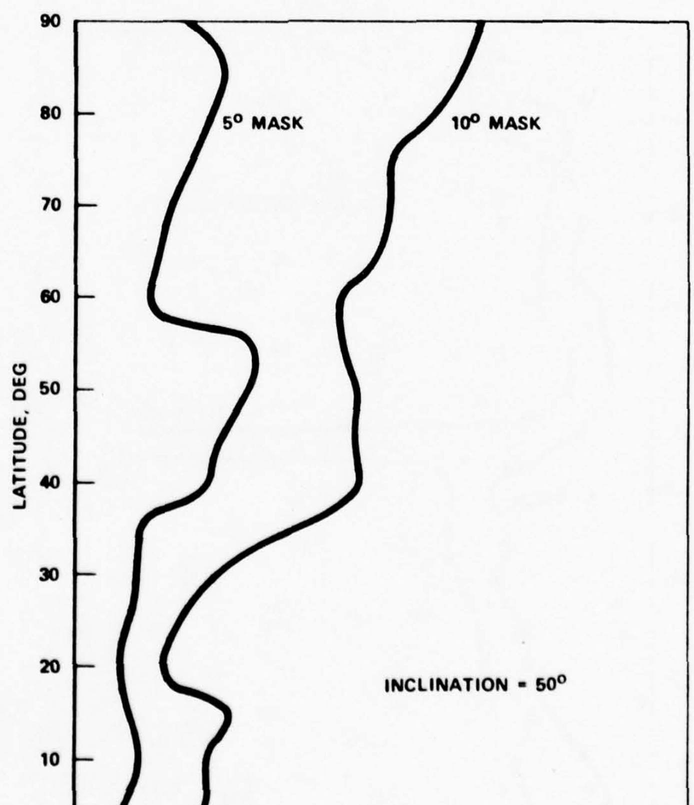


Fig.10 4 x 6 Global latitude performance

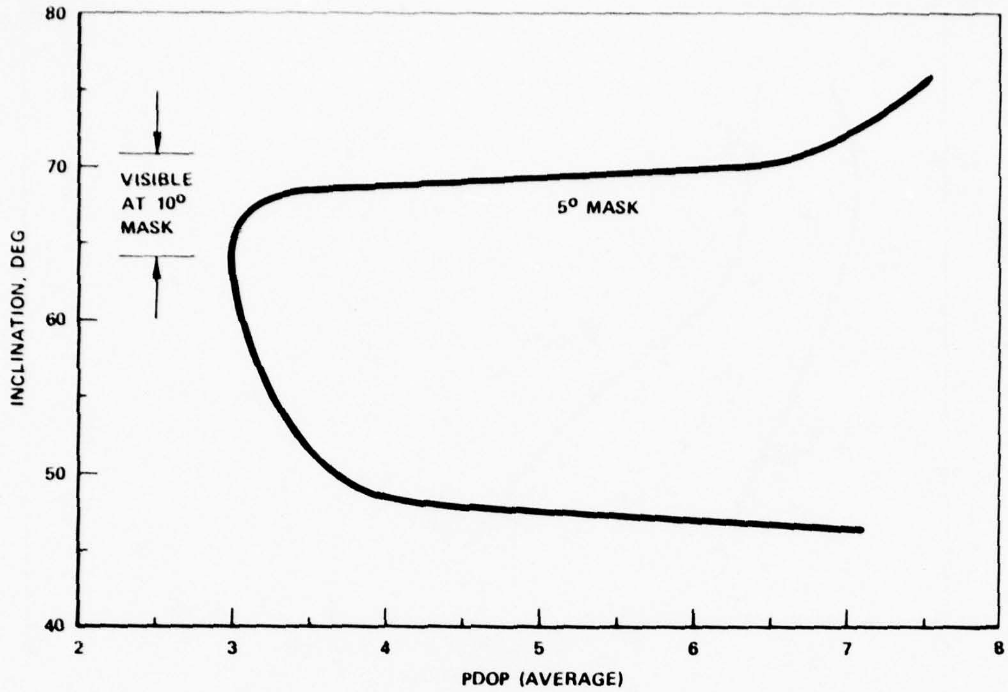


Fig.11 5 x 4 Global geometric performance

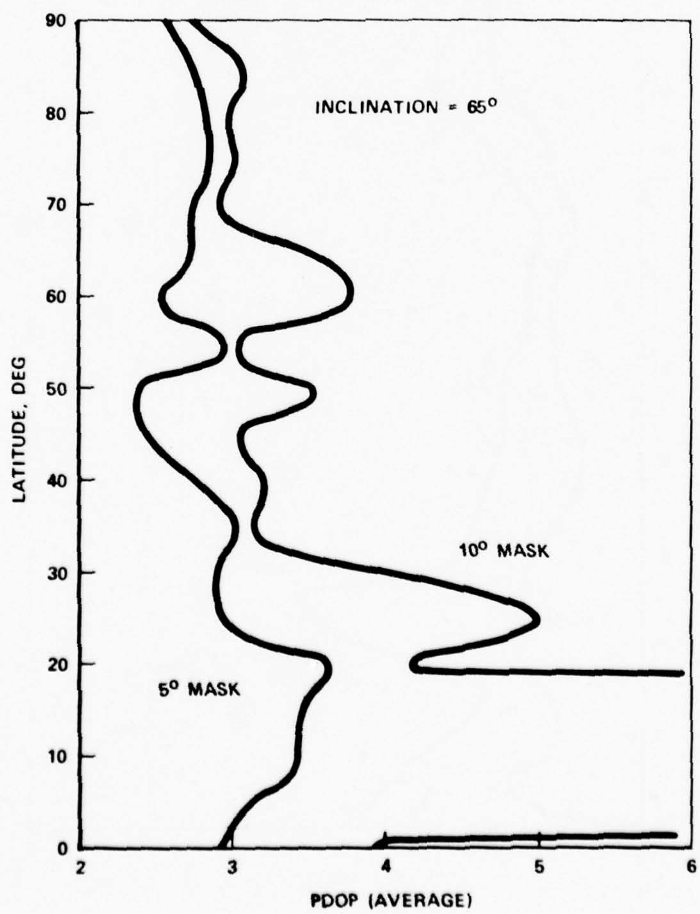


Fig.12 5 x 4 Global latitude performance

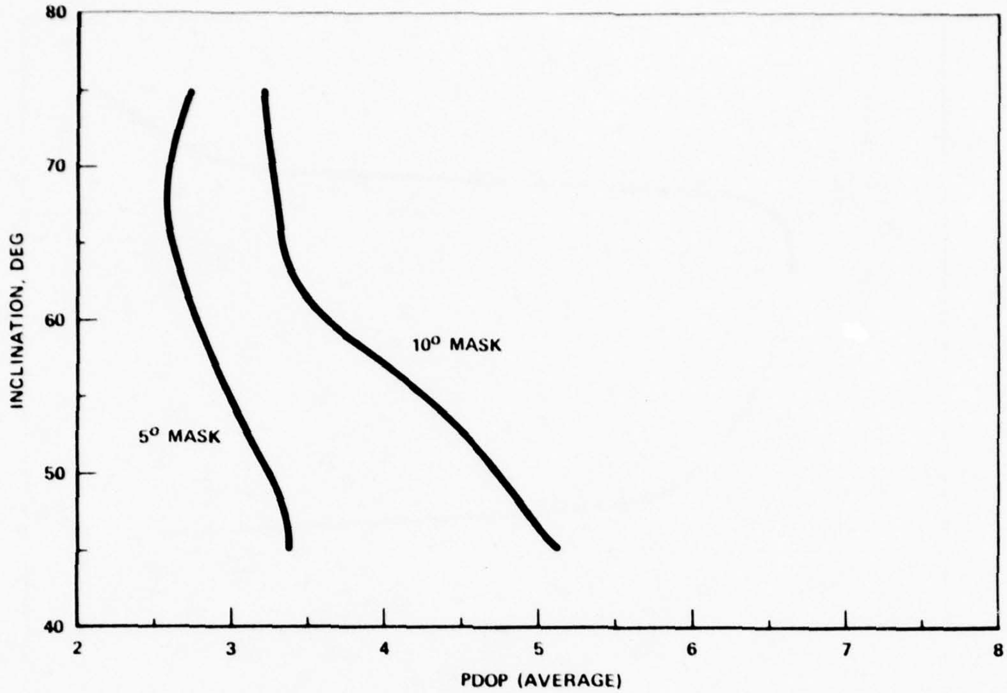


Fig.13 5 x 5 Global geometric performance

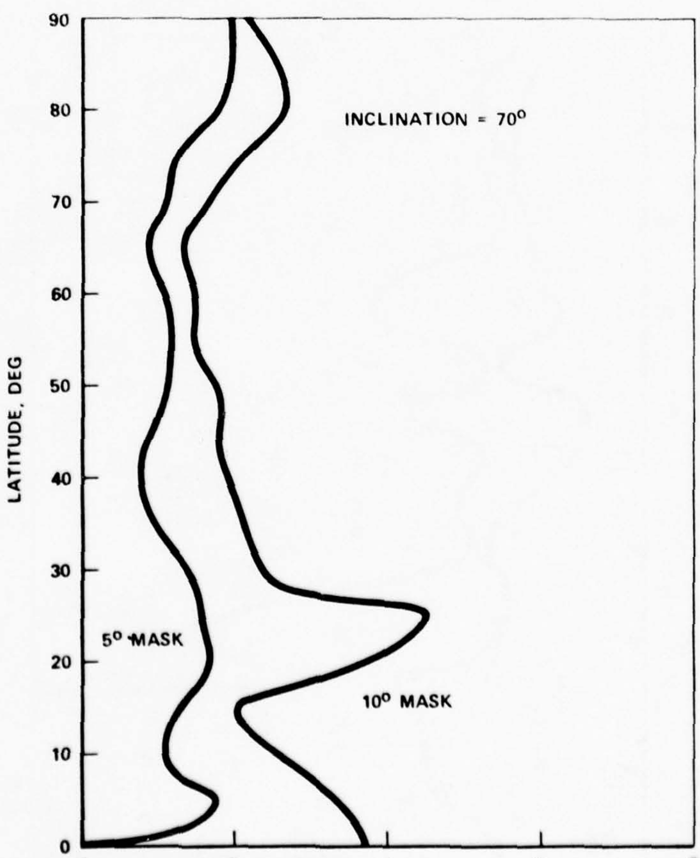


Fig.14 5 x 5 Global latitude performance

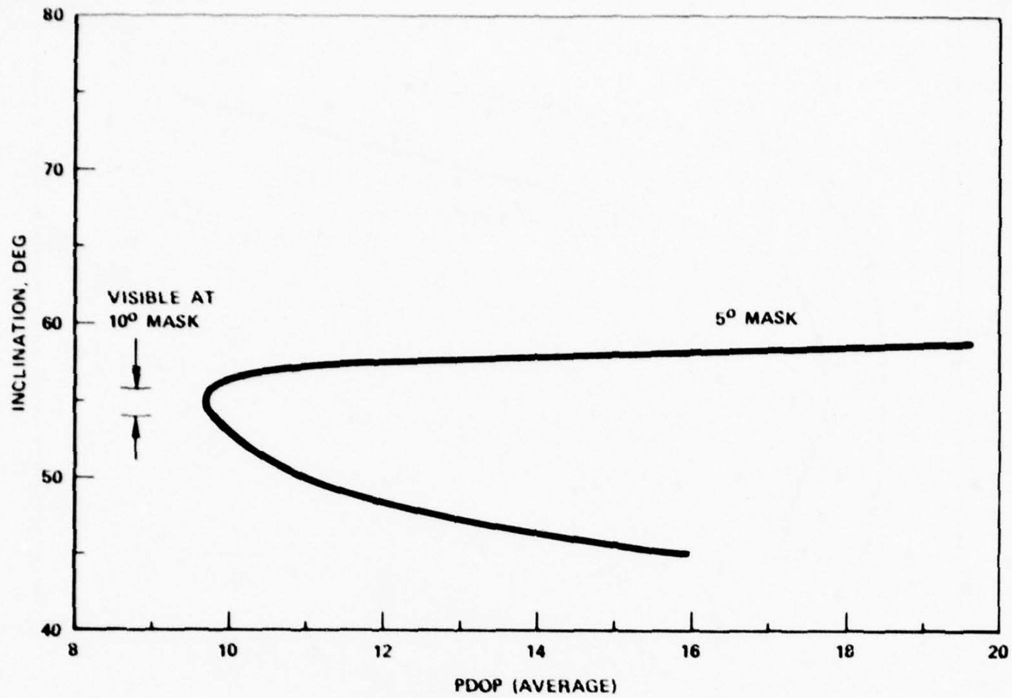


Fig.15 6 x 3 Global geometric performance

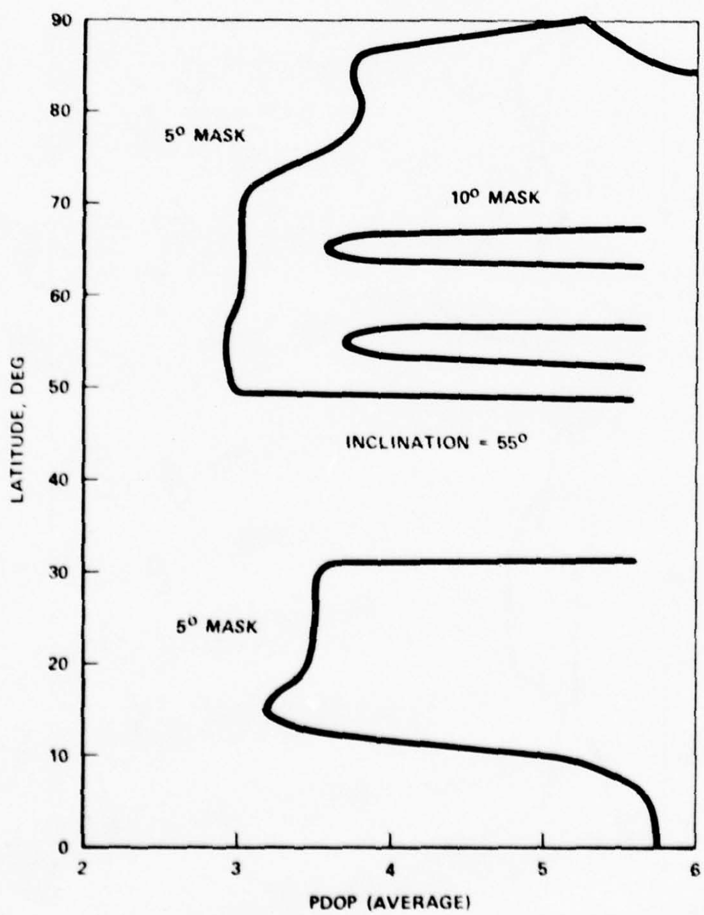


Fig.16 6 x 3 Global latitude performance

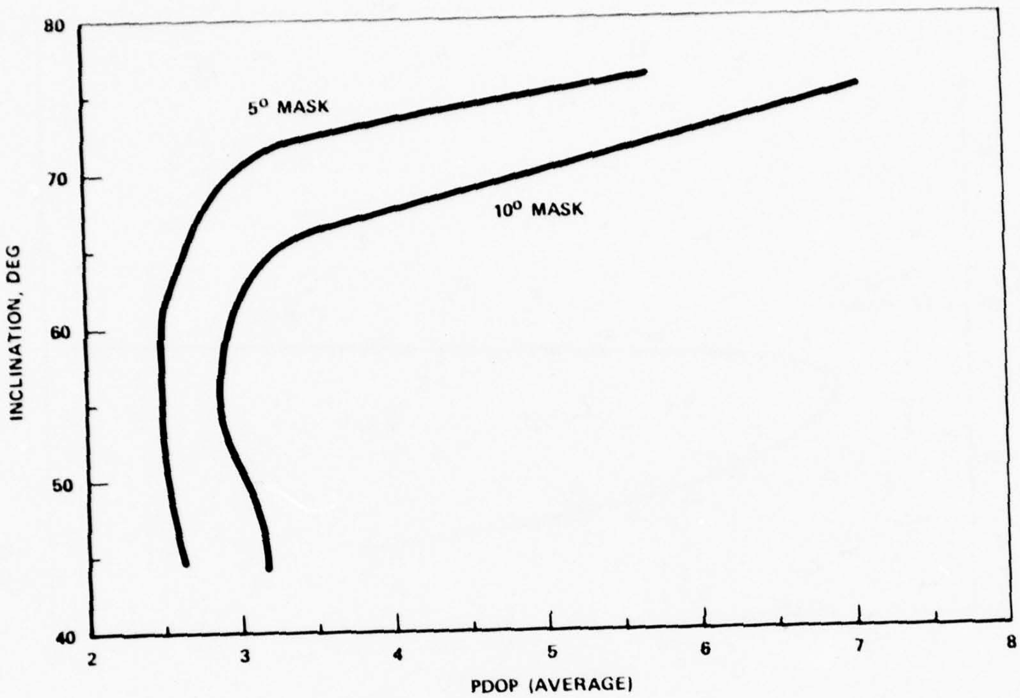


Fig.17 6 x 4 Global geometric performance

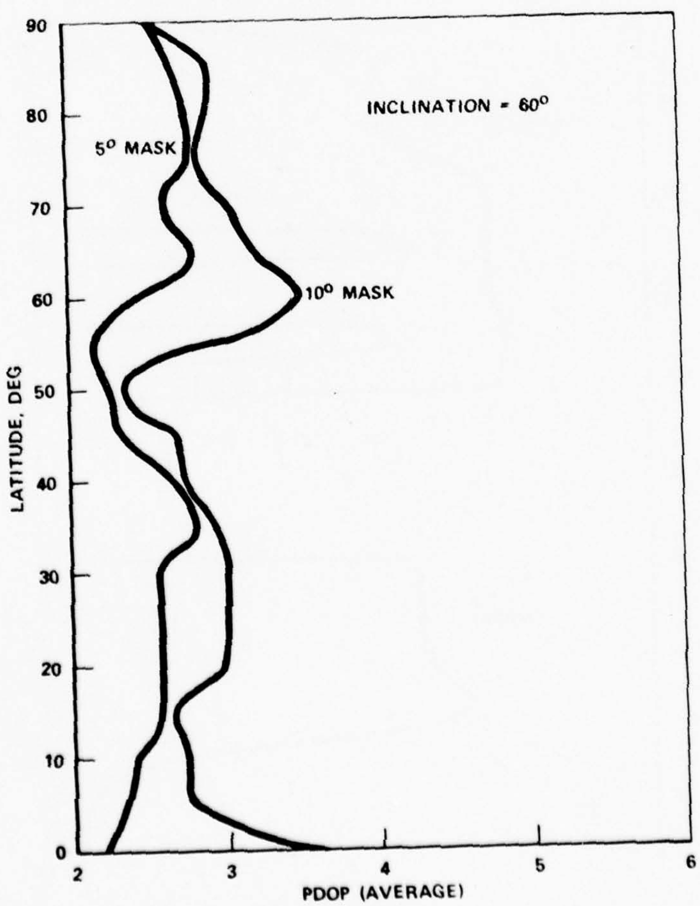


Fig.18 6 x 4 Global latitude performance

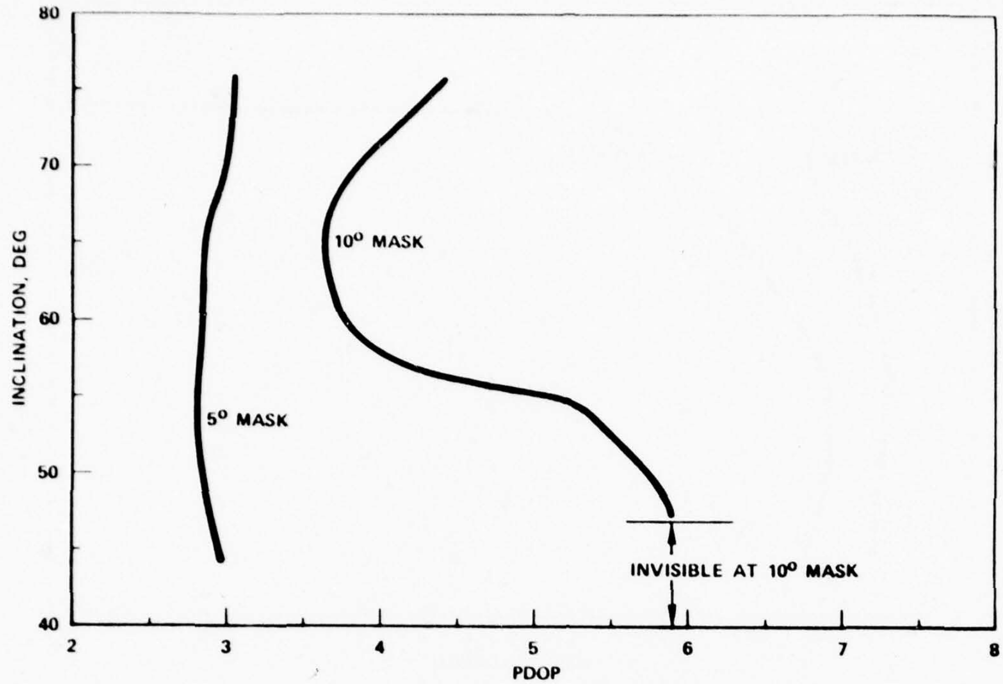


Fig.19 7 x 3 Global geometric performance

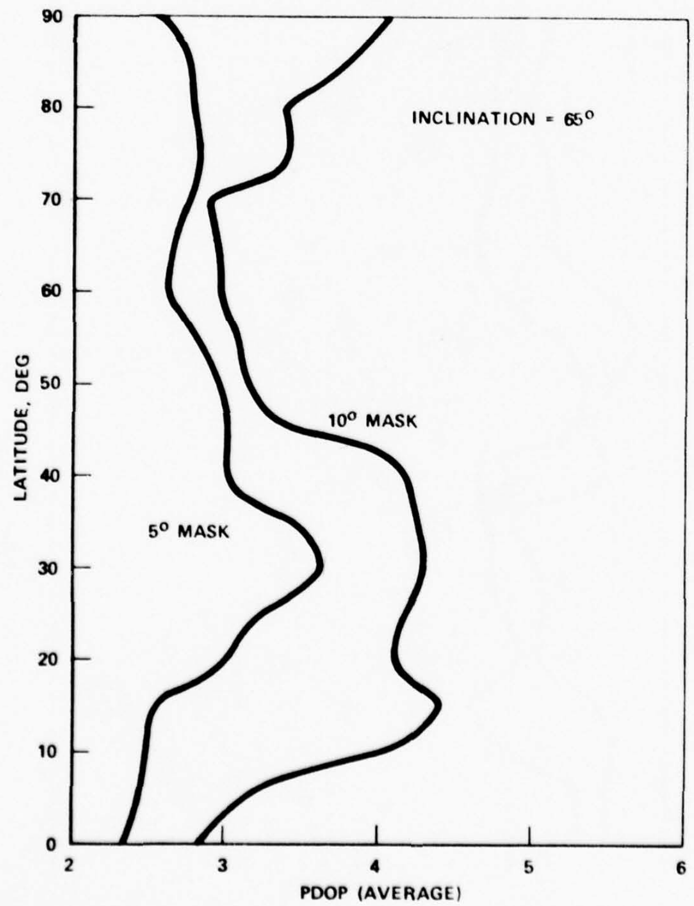


Fig.20 7 x 3 Global latitude performance

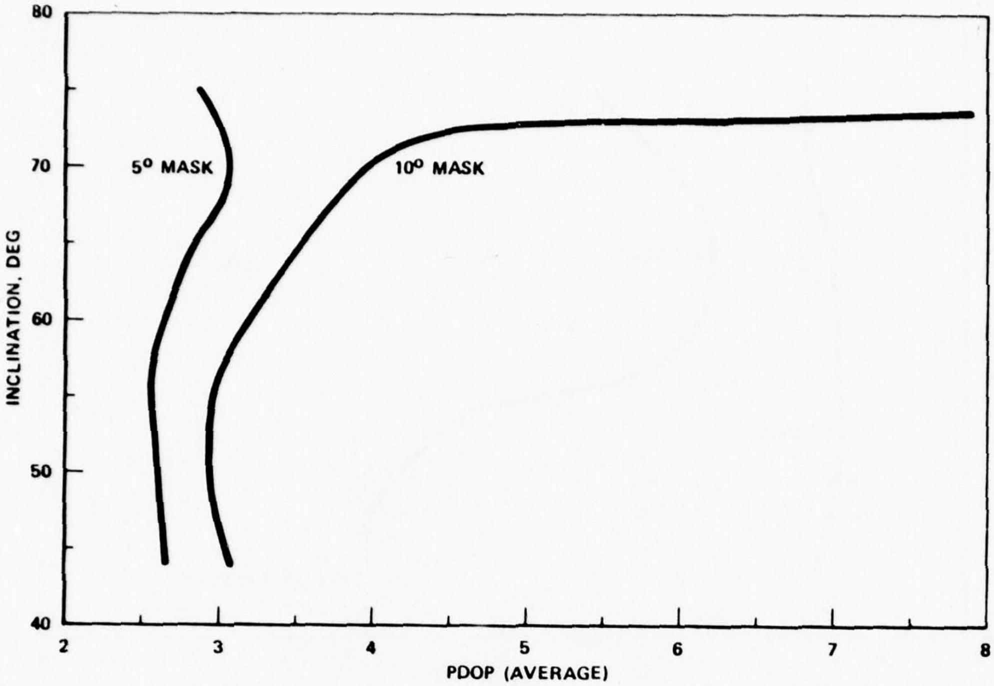


Fig.21 7 x 4 Global geometric performance

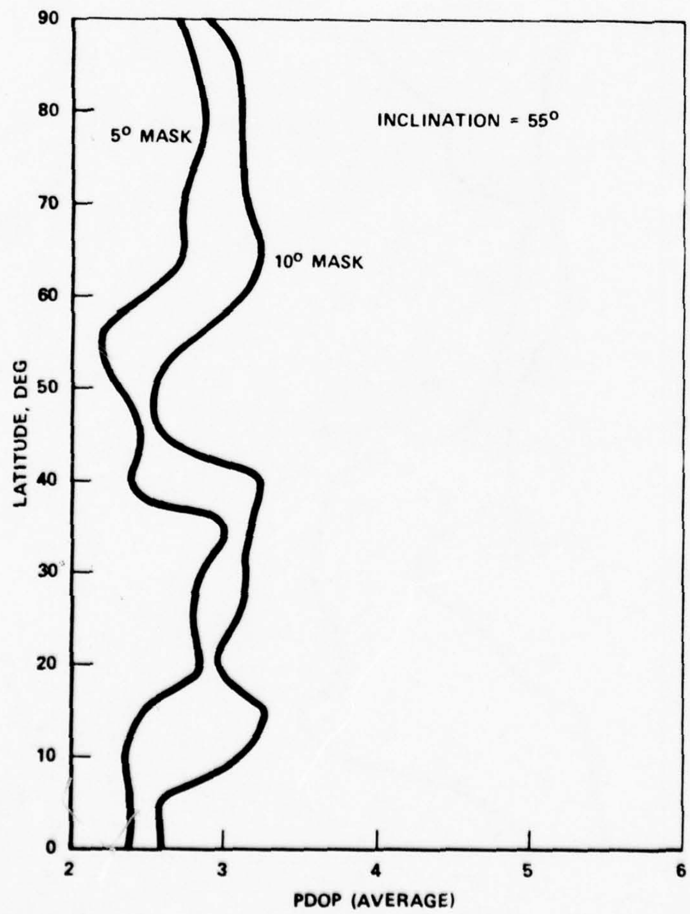


Fig.22 7 x 4 Global latitude performance

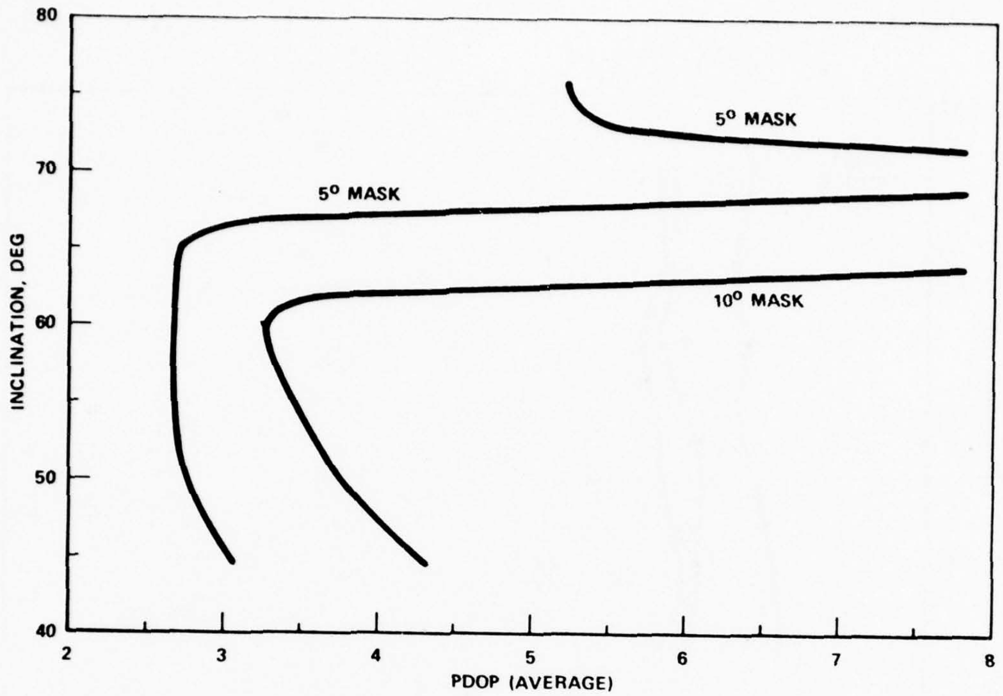


Fig.23 8 x 3 Global geometric performance

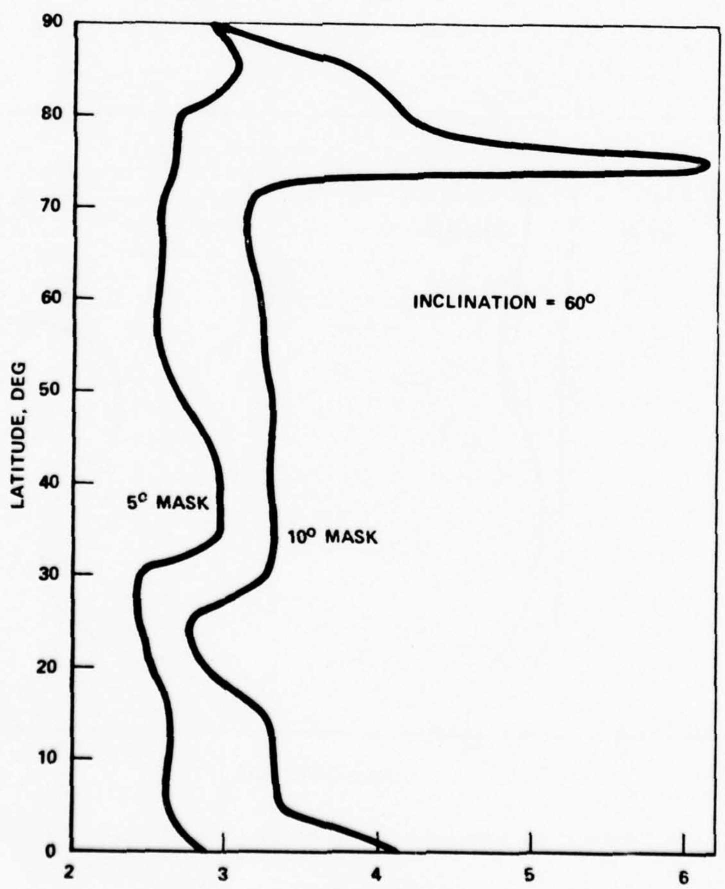


Fig.24 8 x 3 Global latitude performance

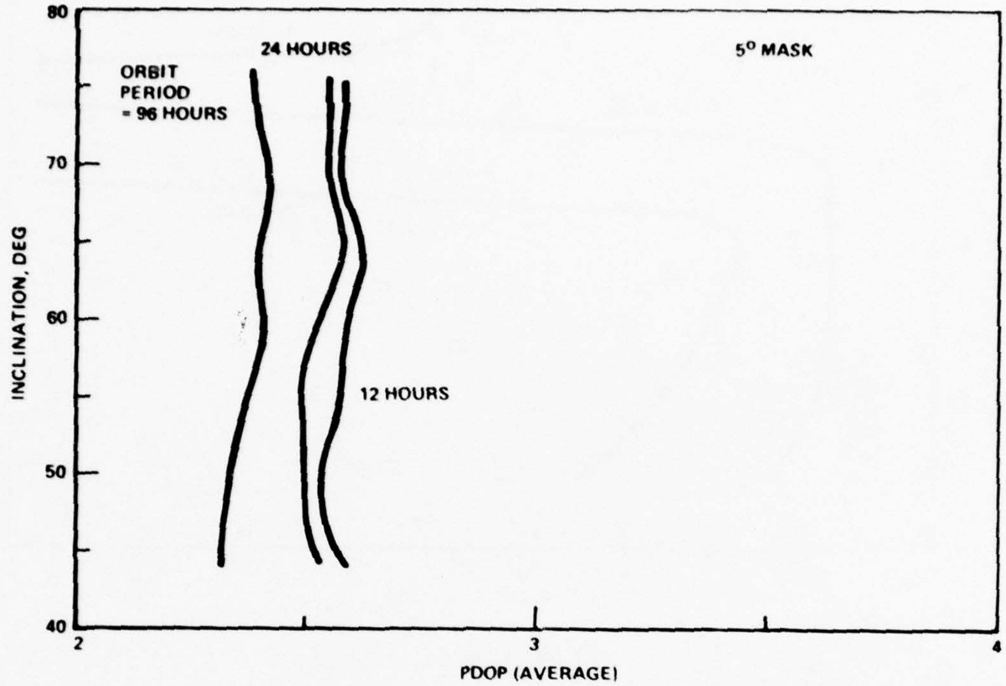


Fig.25 3 x 8 Global geometric performance (5° mask)

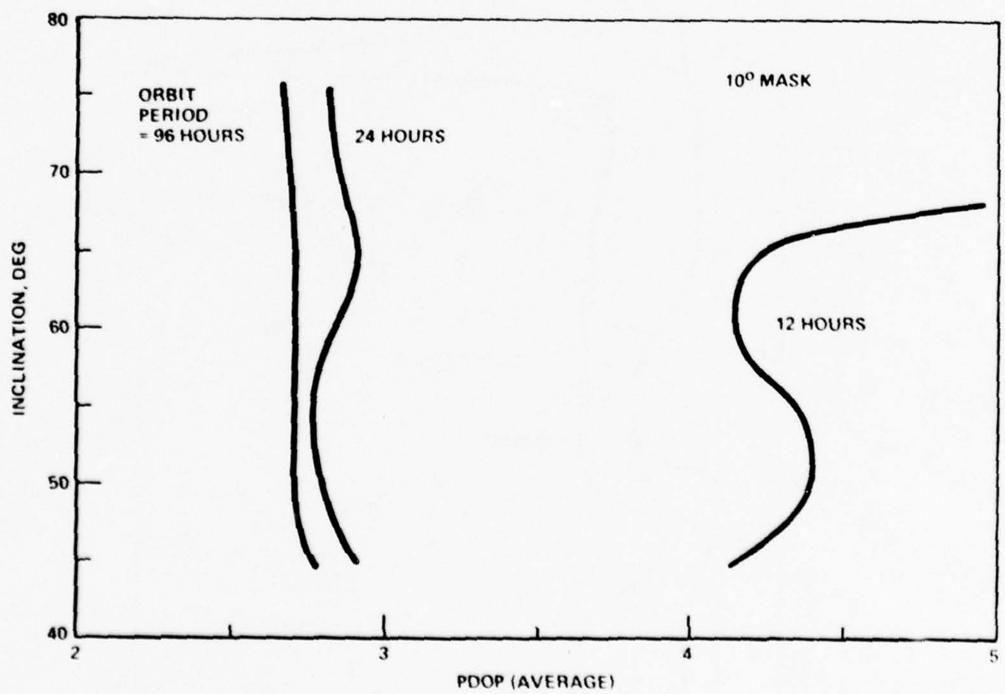


Fig.26 3 x 8 Global geometric performance (10° mask)

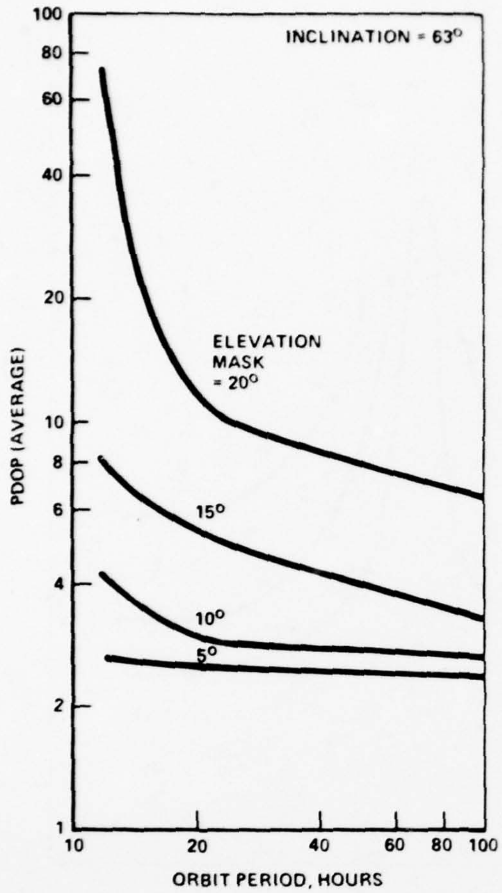


Fig.27 3 x 8 Global orbit period performance

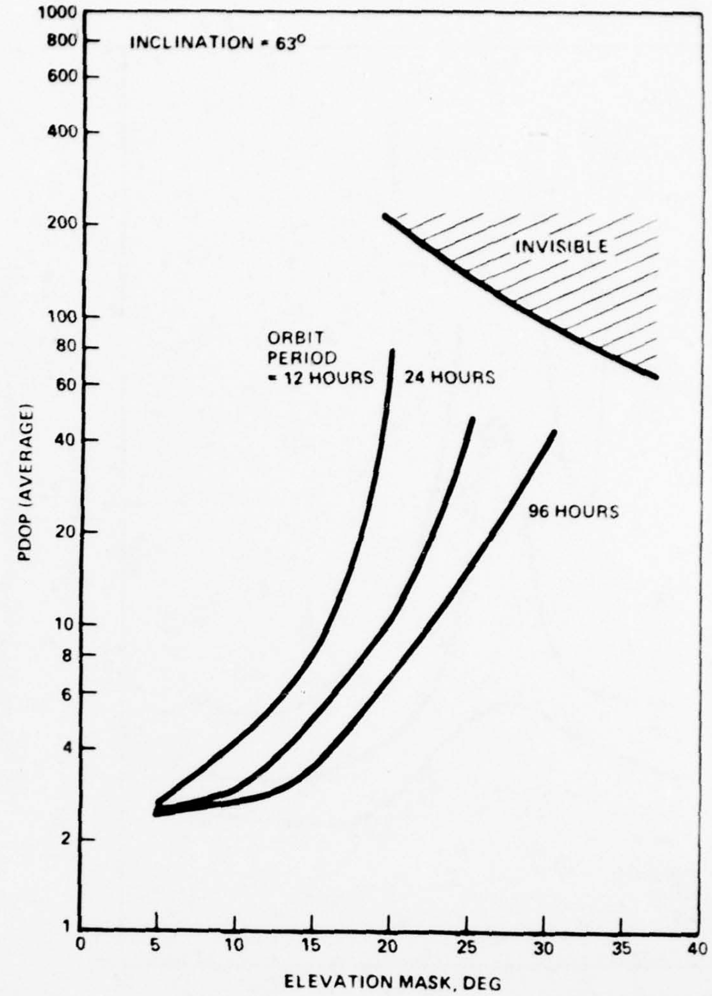


Fig.28 3 x 8 Global elevation mask performance

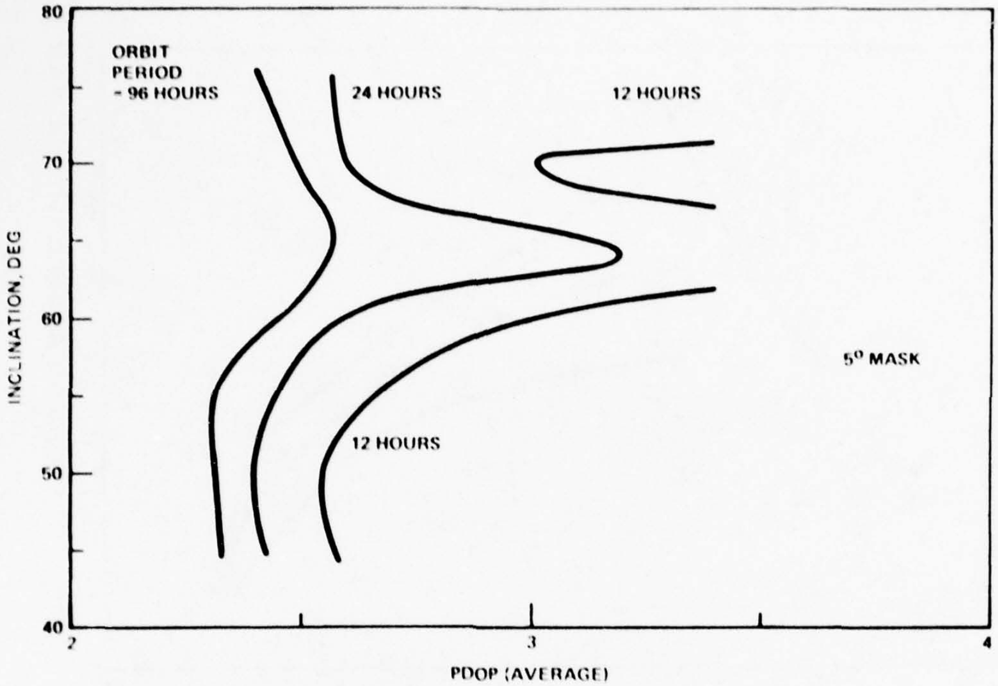


Fig.29 4 x 6 Global geometric performance (5° mask)

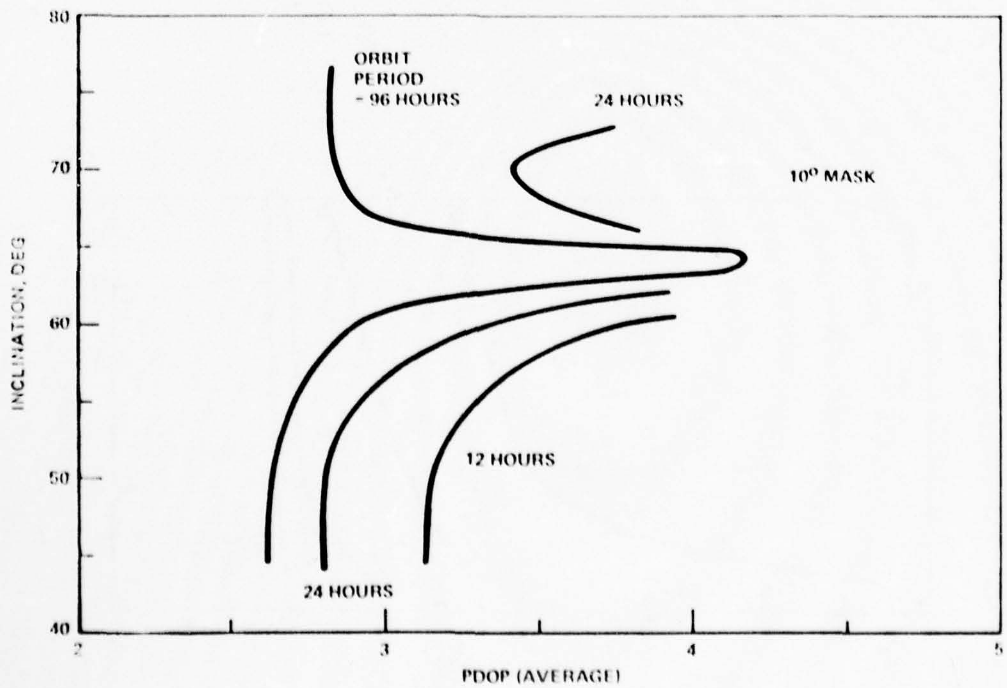


Fig.30 4 x 6 Global geometric performance (10° mask)

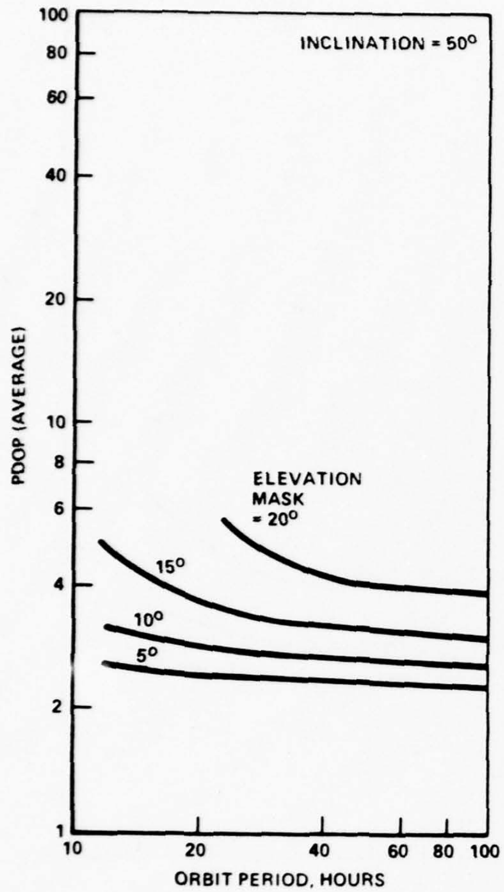


Fig.31 4 x 6 Global orbit period performance

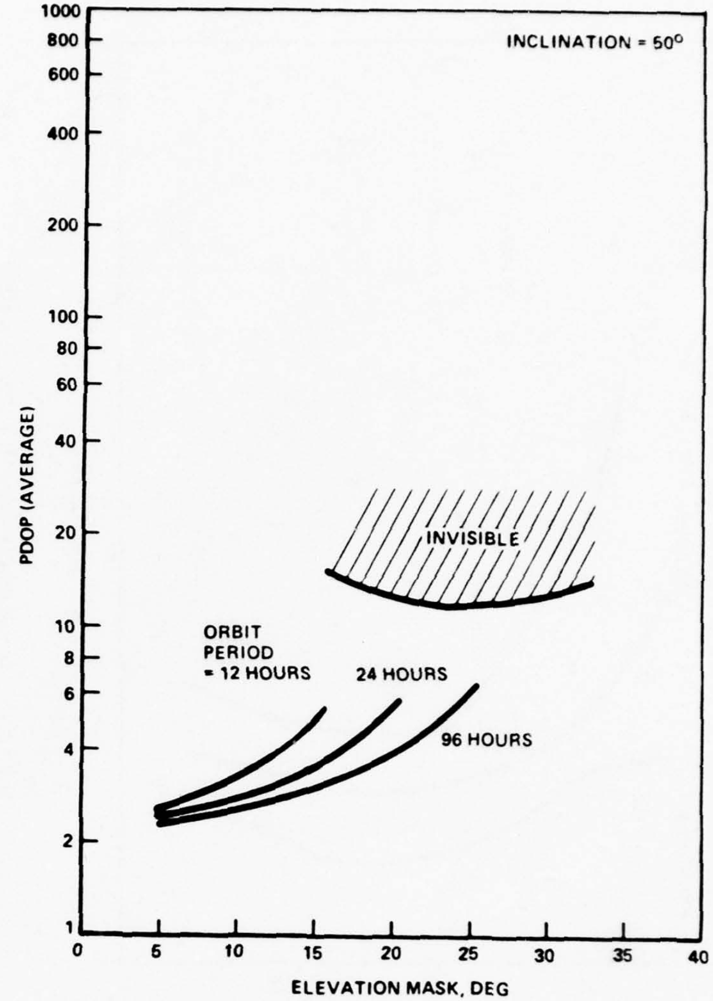


Fig.32 4 x 6 Global elevation mask performance

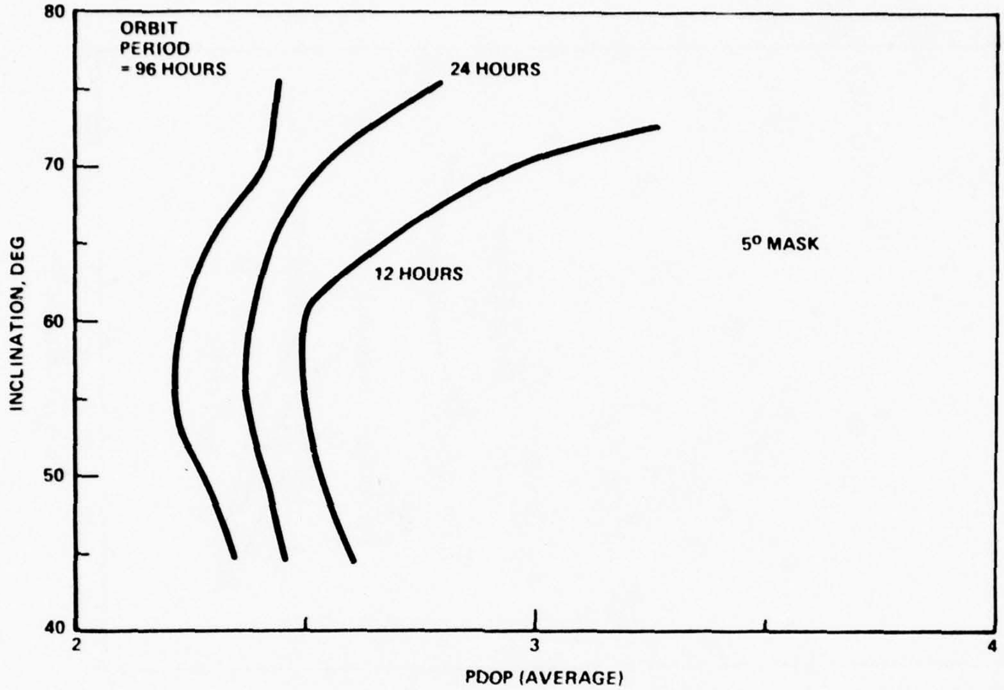


Fig.33 6 x 4 Global geometric performance (5° mask)

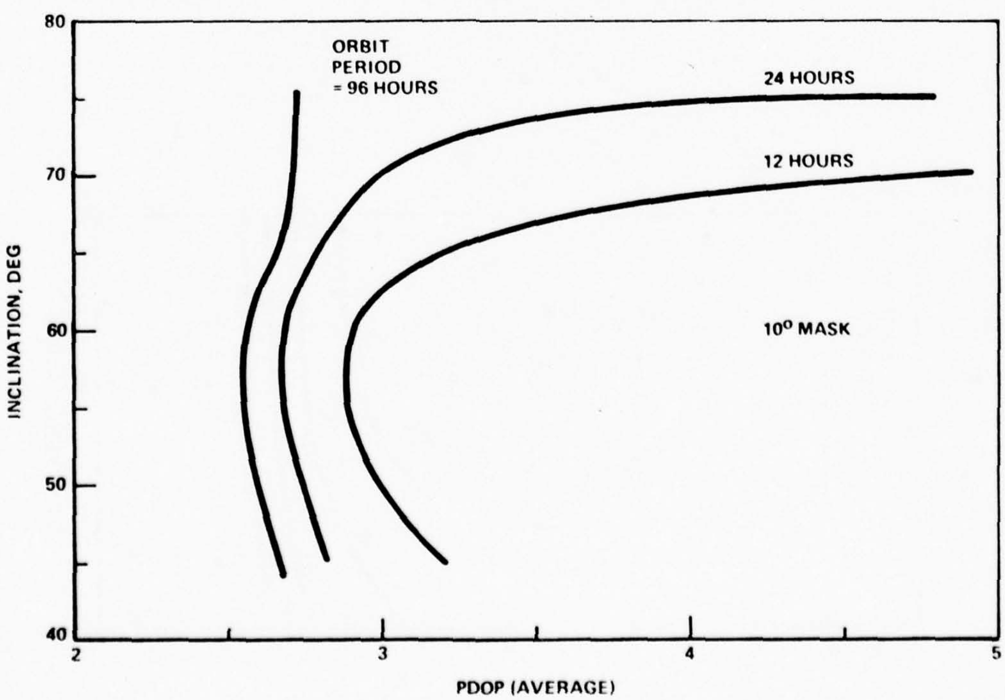


Fig.34 6 x 4 Global geometric performance (10° mask)

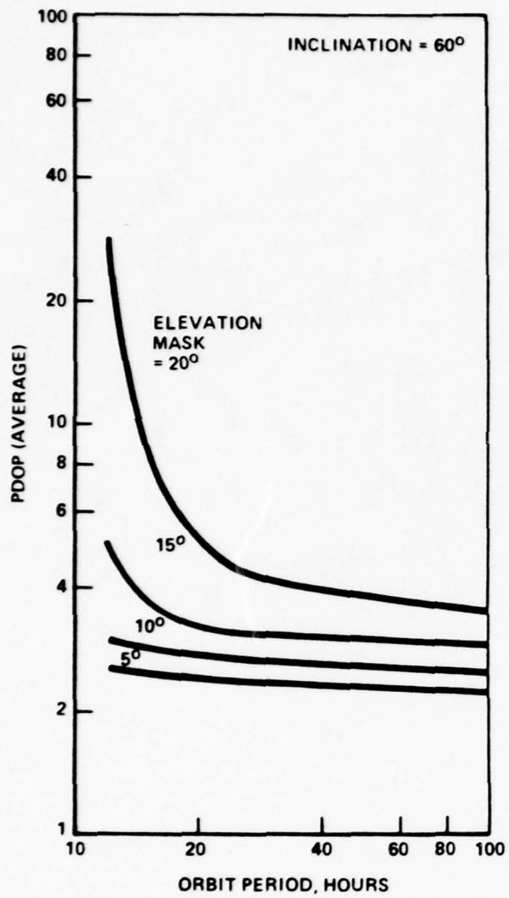


Fig.35 6 x 4 Global orbit period performance

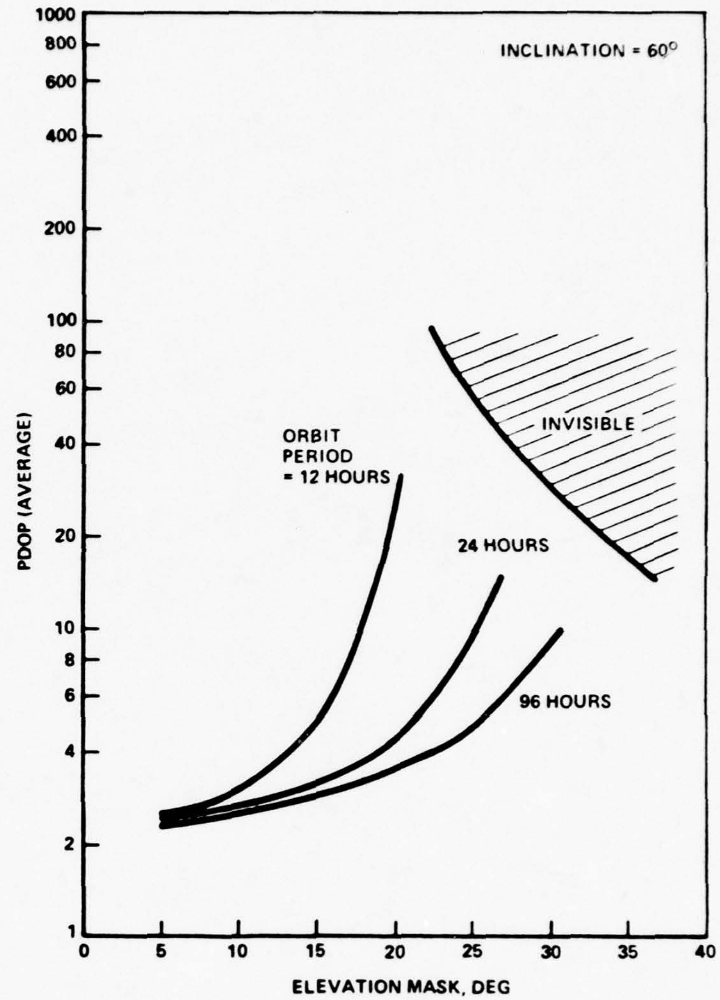


Fig.36 6 x 4 Global elevation mask performance

ON THE OPTIMAL SELECTION OF SATELLITES IN GPS

Katsumi Ohnishi
School of Engineering and Applied Science
University of California
Los Angeles, California 90024
U.S.A.

SUMMARY

The purpose of this paper is to give an anlytical method for the optimal selection of four satellites for the estimation of the user's position in Global Positioning Systems (GPS). It is shown that the present problem can be completely discussed within the framework of linear estimation theory which converts the optimal selection of satellites into the optimal orientation of four antennas. An estimator equation is presented in this paper which has been shown to give better estimation results compared with the estimator equation which has been used for the position determination in GPS up until the present. The four satellites to be selected are obtained by directly minimizing the determinant of the covariance of the estimation error.

1.0 INTRODUCTION

It is well known that four satellites are enough to determine the user's position in GPS (Ref. 1). One of the controversial problems is the selection of four satellites in such a manner that the estimation error of the user's position should be minimized.

Bogen (Ref 2) stated that there was an extremely high correlation between the volume of the tetrahedron and Position Dilution of Prediction (P.D.O.P.). There are two kinds of errors which should be taken into account in the problem of satellites selection, i.e., the errors involved in the user's measurement and those errors in the prediction of satellite position and the satellite clock drift (Ref. 3). It will be shown that these errors should be treated separately.

In this paper, the problem of optimal selection of four satellites is discussed from the point of linear estimation theory. A new estimation equation is presented for the optimal selection, which is shown in Ref. 4 to give better estimation in the case where the two kinds of errors are involved as in GPS. The optimal selection of four satellites in GPS is determined by directly minimizing the determinant of the covariance of the estimation error. It is shown that the error covariances which are involved in GPS uniquely determine the four satellites in an optimal fashion.

The organization of this paper is as follows. In Section 2.0, the problem is mathematically formulated and a solution to the satellite selection is given in Section 3.0. The previous results on the optimal satellite selection (Ref. 1, 2) are discussed from the point of the present method in Section 4.0.

2.0 PROBLEM STATEMENT

According to Ref. 1, we have

$$\begin{bmatrix} e_1' & -1 \\ e_2' & -1 \\ e_3' & -1 \\ e_4' & -1 \end{bmatrix} \begin{bmatrix} R_o \\ B_o \end{bmatrix} = \begin{bmatrix} e_1' \; 1 & & & \\ & e_1' \; 1 & & \\ & & e_3' \; 1 & \\ & & & e_4' \; 1 \end{bmatrix} \begin{bmatrix} R_1 \\ B_1 \\ R_2 \\ B_2 \\ R_3 \\ B_3 \\ R_4 \\ B_4 \end{bmatrix} - \begin{bmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \rho_4 \end{bmatrix} \qquad (1)$$

where

R_o: Position vector of observer

R_i: Position vector of i-th satellite

B_o: Bias caused by the difference between the observer clock and system time.

B_i: Bias caused by the difference between the satellite clock time and system.

$D_i \triangleq R_i - R_o$

$e_i \triangleq D_i/||D_i||$

$\rho_i \triangleq D_i + B_i + B_o$

and z' shall denote the transpose of z throughout the rest of the paper.

Define

$$H \triangleq \begin{bmatrix} e_1' & -1 \\ e_2' & -1 \\ e_3' & -1 \\ e_4 & -1 \end{bmatrix}$$

Considering the case where the measurement has an error, Eq. (1) is written as

$$y = Hx + \varepsilon \tag{2}$$

where $y \in \mathbb{R}^m$, $x \in \mathbb{R}^n$ and ε is the measurement error.

Roughly speaking, the user in GPS can determine his position by measuring the relative position to the satellites whose exact positions are not known and must be predicted with unneglectable errors (Ref.3). Thus, even in the ideal case where the user can make a measurement without errors, he can only determine his position with some errors. Therefore, x in Eq. (2) must be regarded as a random variable which is independent of ε.

Without loss of generality, we may assume the following:

$$\left.\begin{aligned} E\{\varepsilon\} &= 0 \\ E\{\varepsilon\varepsilon'\} &= Q \\ E\{x\} &= 0 \\ E\{xx'\} &= V \end{aligned}\right\} \tag{3}$$

where

Q and V are positive-definite matrices. The matrix Q is the error covariance of the user's measurement while the matrix V is the error covariance of x caused through the prediction errors of the satellite position and the satellite clock drift.

The problem to be considered is stated as follows: Given error covariances Q and V, determine the satellites (or equivalently the matrix H) in order to minimize the estimation error of x .

3.0 OPTIMAL SELECTION OF FOUR SATELLITES

Ref. 4 shows that the linear estimate $\hat{x}$ of x which minimizes $E\{||x - \hat{x}||^2\}$ is given by

$$\hat{x} = (H'Q^{-1}H + V^{-1})^{-1} H'Q^{-1}y \tag{4}$$

with corresponding error covariance

$$E\{(x-\hat{x})(x-\hat{x})'\} = (H'Q^{-1}H + V^{-1})^{-1} \tag{5}$$

Since the estimate given by Eq. (4) depends on H, we shall denote $\hat{x}$ by $\hat{x}_H$, i.e.

$$\hat{x}_H = (H'Q^{-1}H + V^{-1})^{-1} H'Q^{-1}y \tag{6}$$

The problem of determing the optimal selection of satellites can be restated as:

Determine H so that

$\det\,(E\{(x-\hat{x}_H)(x-\hat{x}_H)'\})$ is minimized.

By Eq. (5), we have

$$\det\,(E\{(x-\hat{x}_H)(x-\hat{x}_H)'\}) = (\det\,(H'Q^{-1}H + V^{-1}))^{-1} \tag{7}$$

In GPS we can set H (Ref. 1) as

$$H = \begin{bmatrix} \sin\Theta_1 \cos\psi_1 & \sin\Theta_1 \sin\psi_1 & \cos\Theta_1 & -1 \\ \sin\Theta_2 \cos\psi_2 & \sin\Theta_2 \sin\psi_2 & \cos\Theta_2 & -1 \\ \sin\Theta_3 \cos\psi_3 & \sin\Theta_3 \sin\psi_3 & \cos\Theta_3 & -1 \\ \sin\Theta_4 \cos\psi_4 & \sin\Theta_4 \sin\psi_4 & \cos\Theta_4 & -1 \end{bmatrix} \tag{8}$$

The JTIDS relative navigation technique, as professed in the Hippocratic principle that the treatment shall not leave the patient worse, takes care not to degrade the performance of any members of the community. Recognizing that the tactical grid position may only be loosely related to

This is due to the fact that

$$e_i' = [\sin\Theta_i \cos\psi_i \quad \sin\Theta_i \sin\psi_i \quad \cos\Theta_i \quad -1]$$

where $i = 1, 2, 3, 4$.

In other words, the selection of satellites is equivalent to the orientation of the vectors $e_i (i=1, 2, 3, 4)$, which is equivalent to the determination of Θ_i and ψ_i $(i = 1, 2, 3, 4)$.

A little insight into the practical situations indicates that the orientation of the vectors e_i should be symmetric from each other and information should be obtained from all the directions as evenly as possible. Thus we may set

$$\begin{cases} \Theta_i = \Theta \\ \psi_i = (i-1)\pi/2 \end{cases} \quad (i = 1, 2, 3, 4) \tag{9}$$

where Θ is some constant and $0 \leq \Theta \leq \pi/2$

By Eqs. (8) and (9), the matrix H explicitly depends on Θ, which we denote by $H(\Theta)$. Substituting Eq. (9) for Eq. (8), we have

$$H(\Theta) = \begin{bmatrix} \sin\Theta & 0 & \cos\Theta & -1 \\ 0 & \sin\Theta & \cos\Theta & -1 \\ -\sin\Theta & 0 & \cos\Theta & -1 \\ 0 & -\sin\Theta & \cos\Theta & -1 \end{bmatrix} \tag{10}$$

We assume that

$$\begin{cases} Q = qI_1 \\ V = \delta I_2 \end{cases} \tag{11}$$

where I_1 and I_2 are, respectively, m x m and n x n identity matrices and q and δ are positive constants.

From Eqs. (7) and (10), the problem under consideration is reduced to

$$\max_{0\leq\Theta\leq\pi/2} \det(H'(\Theta)Q^{-1}H(\Theta) + V^{-1}) \tag{12}$$

Eqs. (10) and (11) give

$$\det(H'(\Theta)Q^{-1}H(\Theta) + V^{-1}) = (-2q^{-1}t + 2q^{-1} + \delta^{-1})^2 [(16q^{-2} + 4q^{-1}\delta^{-1} - 16)t + \delta^{-1}(4q^{-1} + \delta^{-1})] \tag{13}$$

where

$$t \triangleq \cos^2\Theta \tag{14}$$

Let f(t) be the right hand side of Eq. (13), then the optimization problem to be solved is:

$$\max_{0\leq t\leq 1} f(t) \ .$$

An easy calculation leads to the folloiwng solutions of the maximizing problem.

If $q^2 \geq 1$ and $\delta > 0$, then $t = 0$ is optimal.

If $q^2 < 1$ and $0 < \delta \leq \frac{2q^3-q}{4(1-q)}$, then $t = 0$ is optimal.

If $q^2 < 1$ and $0 < \frac{2q^3-q}{4(1-q)} < \delta$, then $t = t^*$ is optimal

where

$$t^* = \frac{1}{3} + \frac{2}{3} \cdot \frac{q^3}{4q^2\delta - 4\delta - q} \tag{15}$$

The integrated relative navigation system employs these inputs to perform five basic functions which may be termed

- GRID SYNCHRONIZATION

Recalling the definition given by Eq. (14), we get the following table of the optimal Θ^*.

Case	Conditions	Optimal Θ^*
a)	$q \geq 1,\ \delta > 0$	$\pi/2$
b)	$0 < q < 1,\ 0 < \delta \leq \alpha$	$\pi/2$
c)	$0 < q < 1,\ 0 < \alpha < \delta$	$\cos^{-1}\sqrt{t^*}$

where t^* is given by Eq. (15) and

$$\alpha \triangleq \frac{2q^3-q}{4(1-q^2)} \quad .$$

Since $\Theta^* = \pi/2$ violates the fact that any three vectors out of $\{e_i | i=1, 2, 3, 4\}$ must be linearly independent and the function $f(\cdot)$ is monotone decreasing with respect to t, we must take Θ as close to $\pi/2$ as possible in the cases a) and b) of the above table in the practical situations.

4.0 ON THE PREVIOUS RESULTS

In Ref. 2, ϵ in Eq. (2) is regarded as an error which is due to ephemeris modeling, ionosphere modeling, troposphere modeling, multipath, receiver noise, etc. That is, the errors involved in the problem of satellite selection are not classified as it was done in this paper.

The following estimation equation is presented in Ref. 2.

$$\hat{x} = (H'Q^{-1}H)^{-1}\, H'Q^{-1}y \tag{16}$$

with error covariance given by

$$(H'Q^{-1}H)^{-1} \tag{17}$$

Following the method presented in Section 3.0, H can be determined by minimizing

$$\det\ (H'Q^{-1}H)^{-1} \tag{18}$$

or equivalently by maximizing

$$\det\ (H'Q^{-1}H) \tag{19}$$

Theorem

If H is singular and Q is positive definite, then there does not exist matrix H which gives physically reasonable estimate of x.

Proof

It is enough to show

$$\det\ (H'Q^{-1}H) = 0$$

Since H is singular, the null-space $N(H)$ of H contains a non-zero vector, say u, i.e.,

$$Hu = 0$$

and

$$u \neq 0$$

Then clearly

$$H'Q^{-1}Hu = 0$$

and

$$N(H'Q^{-1}H) \neq (0)$$

Thus $H'Q^{-1}H$ is singular

Q.E.D.

The Locator Routine updates stored coordinates using geographic and grid position information, sensor detected range, bearing, mark-on-top, etc., course and speed estimates from tracking algorithms and other sources. Grid positions thus assigned are transmitted to other CNI routines as required.

***

Theorem indicates that the estimator equation Eq. (16) does not give a reasonable estimate of x in the case where H is given by Eq. (10).

5.0 CONCLUSIONS

An analytical method to the problem of the selection of four satellites is given which shows the problem can be discussed as one of the linear estimation problems. The result indicates that the noise covariances defined by Eq. (11) uniquely determine the four satellites which give the minimum value of determinant of the estimation error covariance.

REFERENCES

1. Yonezawa, K., "Geometric Performance of Global Positioning System," F04701-75-C-0211, Mod. 1, U.S. Air Force Space & Missile Systems Organization, 1975.

2. Bogen, A.H., "Geometric Performance of the Global Positioning System," TR-0074(4461-02)-1, Aerospace Corporation, June 1974.

3. Schaibly, J.H. et al., "Simulated and Projected Performance of the NAVSTAR GPS Control Segment," AIAA Guidance and Control Conference, San Diego, California, August 16-18, 1976.

4. Leunberger, D.G., Optimization by Vector Space Methods, Chapter 4, John Wiley, 1969.

IV. SYNCHRONIZATION AND RANGING

Robert Stow, Singer-Kearfott Corporation

THE EVOLUTION OF JTIDS

RONALD S. VAUGHN
NAVAL AIR DEVELOPMENT CENTER
SENSORS AND AVIONICS TECHNOLOGY DIRECTORATE
WARMINSTER, PA 18974

SUMMARY

The Joint Tactical Information Distribution System (JTIDS) evolved from separate developmental programs initiated independently by the US Air Force and the US Navy. The genesis of these predecessor programs is traced and the reasons for merging them into a joint services development program. In addition, the major technology evolution affecting the JTIDS engineering approach is touched upon.

The Joint Tactical Information Distribution System (JTIDS) is a full scale tactical command and control system under joint development by all the Armed Forces of the United States under the executive leadership of the US Air Force. JTIDS has planned applications throughout a broad range of scenarios for the various services and embodies the most advanced technology of any tactical system of such broad scope currently planned. The following papers expose various aspects of the system concept, architecture, engineering applications, and intended use of the system.

It would be historically inaccurate to leave the impression with readers of this tome that the JTIDS concept rose in full form like Aphrodite from the sea without the labor pangs of dedicated engineers, technologists, and service officers. This introductory chapter is meant to limn some major episodes and influences in the evolution of the JTIDS concept as viewed by one present and contributing during the period of conception, and an interested observer at the birth of JTIDS. This chronicle is intended to provide recognition to the contributions made by numerous individuals, unnamed herein, who contributed through their respective agencies to the evolution of the JTIDS concept.

JTIDS did not originate the basic technologies which it utilizes, but, as with many successful engineering programs, it synthesized numerous emerging key technologies into a specific system concept and provided impetus for some specific theoretical concepts to be reduced to practical engineering concepts.

The emergence of digital technology coincided with the transition of command and control systems from a reliance on continuous wave (CW) technology to pulse technology. For example, precision radio navigation technology of the 1940's and 1950's relied heavily on phase measurements of signals from closely controlled, coordinated, essentially CW transmissions from "master" stations (e.g. Loran, Omega Systems). Such systems were limited partly by the difficulty in providing each user with a sufficiently accurate self-contained clock to permit direct comparisons of phase or time-of-arrival (TOA) information, so that such navigation systems were limited to comparisons of phase between receptions from several such master stations. In early digital communication systems, however, the relative clock inaccuracies were overcome by providing a synchronization signal from the transmitter on which the receiver system would lock, effectively eliminating the possibility of direct phase or TOA measurements. Such systems were quite naturally separate and distinct and there was little reason to combine communication and navigation systems into common hardware. The development of the accurate (1 ppm) local digital clocksin a form accessible to every user was a key technological step toward JTIDS.

The essentially pulse waveform technology base of JTIDS was provided in large part by developments such as TACAN/DME. These systems introduced cooperative pulse transmission technology on a large scale in the 1950's and were the first systems to treat extensively the problems of multiple users in a limited spectrum, albeit in an uncoordinated way. As one of the first pulse oriented systems with broad applications, the TACAN system made an easy, natural transition from analog electron tube technology of the late 1950's to digital, solid state technology of the mid-to-late 1960's.

The multiple user, random access characteristics of the TACAN/DME system lead to the development of a system which was extremely conservatively designed for its intended use. This conservatism was key to the eventual selection of the L_X band for the application of JTIDS (See Chapter IIID). The resulting underutilization of this portion of the spectrum by TACAN/DME service led to several exploratory studies by the Naval Air Systems Command in the mid-1960's which considered the feasibility of utilizing the TACAN pulse transmission characteristic as the basis for a secure tactical communications service in addition to the navigation service. These studies can reasonably be considered to be the genesis of the integrated communication-navigation-identification (ICNI) functions later embodied in JTIDS.

The RTT sync mode makes use of a digital filtering technique for correcting relative clock and oscillator errors within the TDMA transceiver. The technique employs a linear filter with time varying gains to estimate relative clock and oscillator frequency errors from measurements of community time relative to the RTT interrogated unit. This iterative clock estimation procedure allows rapid time synchronization to accuracies which will allow accurate interpretati

Several efforts were begun in the late 1960's by the US Air Force and US Navy which directly influenced the course which was to result in the JTIDS program. These programs were severally conducted under exploratory development programs of these services (research and development efforts in command and control). In their earliest phases, these programs were called Integrated Tactical Navigation System (ITNS), Integrated Tactical Air Control System (ITACS), both pursued by the Naval Air Systems Command, and Seek Bus, sponsored by the USAF Electronics Systems Division.

The ITNS program arose from several influences. In the late 1960's the Naval Air development Center was investigating a broad range of synergistic combinations of navigation techniques and equipments under the Advanced Optimal Navigation Equipment (A-ONE) program sponsored by the Naval Air Systems Command. At the same time, advanced satellite navigation programs eventually leading to the NAVSTAR Global Positioning System (GPS) were under study. As a low cost alternative to potentially expensive satellite navigation, the Chief of Naval Development directed the A-ONE program to concentrate on a specific tactical combination of inertial navigation and emerging TOA technology in a concept pioneered by the Singer Company. This concept, expanded and further developed by the Naval Air Development Center, became the ITNS concept which is the heart of the navigation portion of JTIDS. The key technologies utilized by ITNS were the combination of emerging TOA technology and inertial technology with a novel application of Kalman filter theory. Chapter IIIB and its sub-chapters explore this contribution in detail. The ITNS concept was deemed central to open-ocean naval operations of combined naval surface and air forces of the future and was thus explored technically and operationally in various scientific and fleet exercises (see Chapter V. A. 2).

In parallel with the ITNS development, the Naval Air Systems Command was exploring the applications of various radio technologies to airborne communication systems which would be data secure, jam resistant, and would be difficult to intercept or detect. As a part of this ITACS program, the Naval Research Laboratory investigated the application of spread spectrum, frequency hopping, and time hopping in time division multiple access (TDMA) systems. All of these techniques were to find application in ITACS and eventually in JTIDS. Under the leadership of the Naval Air Development Center, these techniques were made a part of a broad ITACS architecture addressing command and control communications across many elements of the radio spectrum, including HF, UHF, and L_X bands.

Since ITACS and ITNS were administered from a single office within the Naval Air Systems Command and both were executed by the Naval Air Development Center, the two programs were closely coordinated. When it was deemed that their individual sub-goals were sufficiently achieved, the two programs were merged in 1973 for closer coordination of detailed hardware development.

Meanwhile, the Air Force Electronic Systems Division had begun the Seek Bus program. The concept engineering for Seek Bus was performed by the Mitre Corporation in response to an emerging requirement from the Tactical Air Command for broader information sharing in tactical command and control, with the Airborne Warning and Control System (AWACS) in the vanguard of this new requirement. Seek Bus could trace its ancestry to a previous Mitre Corporation exploratory development called Position Location, Reporting and Control of Tactical Aircraft (PLRACTA). In the PLRACTA program, elements of electronic position measurement and position reporting were explored with emphasis on jam-protected communications. The Seek Bus program, which emerged in the early 1970's, was more specifically directed to anti-jam communications for the E3A AWACS aircraft mission.

A large degree of technical cooperation and interchange existed between these programs beginning in the late 1960's. As each program in its early phase was directed toward specific non-duplicative but related goals, the Mitre Corporation and the Naval Air Development Center compared technical approaches frequently.

In addition, the Defense Directorate for Research and Engineering (DDR&E) monitored these technical efforts until the confluence of two elements suggested management action. As specific technical solutions were demonstrated by the several programs, specific service applications began to emerge and DDR&E decided that interoperability was an issue that required attention. At the same time, a convergence of the Seek Bus and ITNS/ITACS programs on a similar, but not identical, set of technical approaches indicated that single common solution, or at least a controlled family of solutions, would brake the proliferation of disparate equipments in the defense inventory. In 1973, DDR&E directed the services to establish a joint development program to satisfy the needs of all with anti-jam L-band ICNI system with tactical grid relative navigation capability.

In response, the Joint Logistics Commanders recommended the establishment of JTIDS, a joint services program office (JPO), under the executive leadership of the USAF. All four services have deputy program managers at the program office and all of the technical efforts of the Air Force Seek Bus and the Navy ITNS/ITACS programs* were placed under the cognizance of that office in 1974.

---

* An exception being the most advanced ICNI multifunction architectural development in the other-than-L-band portions of the spectrum. This part of the program is still peculiarly Navy in its formative stage, and is known as the Tactical Information Exchange System (TIES), see Section III. A.

Position References, Navigation Controller and certain other select units called Primaries use round trip timing for clock synchronization. Other units synchronize passively using multiple pseudo-range measurements.

Close continuing technical interface between the Naval Air Development Center and the Mitre Corporation before the merger made the transition relatively simple. The Mitre Corporation had invited the Naval Air Development Center to participate in the selection of the Seek Bus contractor, the Hughes Company (see Section IV. D. (1)), and the Naval Air Development Center had involved the Mitre Corporation in the evaluation of the ITT Corporation concept for distributed time division multiple access to become a part of JTIDS (see Section IV. B).

The JTIDS JPO was able to quickly move forward in two areas: the establishment of a mutually inclusive set of operational requirements (initially for the Navy and Air Force, later to include the Army and Marine Corps) and to establish a mode of interoperability for the several types of equipment being developed under the predecessor programs.

The two services' technical agents were able to quickly settle upon a common signal waveform for JTIDS to provide for interoperability between the two earliest systems already being built by Hughes and Singer (see Sections IV. D. (1). a. and IV. D. (2), a.) and to provide for the growth to the Distributed, Advanced and Hybrid TDMA forms and to insure their interoperability with the earlier systems (see Sections IV. B. and IV. C.). At the same time, the strong merged technical team agreed that a highest risk item in the development program was the as-yet untried Reed-Solomon block error coding and decoding scheme. The Naval Air Development Center immediately contracted with the ITT Corporation and the TRW Corporation to demonstrate this key element of the system design.

The value of the merger of the programs has been clearly demonstrated both in the area of joint requirements and interoperability and in the formation of a very strong merged technical team. The widowm of permitting independent parallel approaches during concept development and merging the best features of all approaches, only when the objectives are clearly common, will be further validated by the multiplicity of successful system types which are about to be demonstrated in flight tests.

Having acquired the relative grid, continual range measurements are used to maintain the grid lock of the User. In addition, the filter is designed to process three other observation types. Each of these is related to the use of geographic information for improved geodetic navigation. They are:

JTIDS SYSTEM OVERVIEW

Roy L. Eisenberg
Technical Director, Engineering Division
ESD/DCBE
JTIDS Joint Program Office
Hanscom AFB, MA 01731 USA

SUMMARY

1.0 BACKGROUND AND SYSTEMS CONCEPT

The JTIDS concept has evolved as a solution to the deficiencies of current line-of-sight tactical communications, navigation and identification systems in the overall military command, control and communications environment. In the past, separate systems of "black boxes" have been developed to meet individual requirements or deficiencies in these areas and the results have been limitations in system capacity and coverage, lack of connectivity between diverse systems, low levels of survivability, obsolescence of data and inability to correlate identification and position data on friends, neutrals and hostiles. JTIDS addresses all of these deficiencies through a development that provides a secure, jam resistant, fully integrated communications, navigation, and identification (ICNI) system. The features which JTIDS integrates are summarized in Figure 1.

FIGURE 1
JTIDS FEATURES

(1) SECURE, JAM-RESISTANT INTEGRATED COMMUNICATIONS, NAVIGATION, IDENTIFICATION (ICNI) FOR TACTICAL COMBAT ENVIRONMENT.

(2) COMMUNICATIONS
- INFORMATION DISTRIBUTION
- DIGITAL VOICE
- COMMON FRAME OF REFERENCE

(3) NAVIGATION
- RELATIVE
- GEODETIC CORRELATION
- TACAN

(4) IDENTIFICATION
- INHERENT
- MARK XII

1.1 Communications Functions

JTIDS provides the critical communications functions required for combat operations in the tactical theatre:

(1) Information distribution of such critical data as position information on friendly participants, track information on hostiles, threat warning and control and vectoring information.

(2) Digital voice for use where the data distribution function will not suffice and a supplementary capability is required.

(3) A common frame of reference provided by the JTIDS relative navigation function so that data can be exchanged in clear, unambiguous position coordinates.

1.2 Navigation Functions

JTIDS also provides the following navigation functions without which the accurate correlation of position data from mutiple data sources would be difficult, if not impossible:

(1) A precise relative navigation capability which provides a common grid and precise relative positioning capability to all participants in a JTIDS network.

(2) Geodetic correlation of the relative grid which provides geodetic alignment and positioning of the relative grid when two or more JTIDS users in a network are either accurately surveyed or can provide other sources of geodetic data such as satellite navigation data to their JTIDS terminals.

(3) Emulation of the TACAN interrogator function in all JTIDS tactical aircraft terminals to provide backward interoperability with current TACAN beacons and elimination of redundant TACAN "black boxes" on JTIDS equipped tactical aircraft.

- The concept of a grid frame frees the relative navigation problem from an absolute frame of reference.
- Precise inter-unit range measurements are the basis for accurate relative positioning schemes.

1.3 Identification Features

The following identification features of JTIDS complete the ICNI capability:

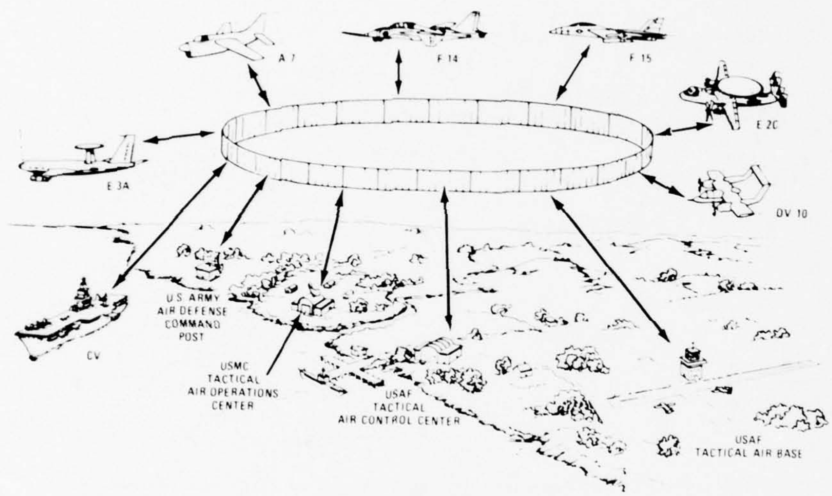
(1) Inherent identification of all JTIDS equipped platforms through interchange of position and identification data by all JTIDS equipped platforms.

(2) Emulation of the Mark XII IFF transponder function in all JTIDS Phase II command and tactical platform terminals for backward interoperability with all Mark X/XII IFF interrogators.

1.4 The JTIDS Network

The basic JTIDS building block is a single communications circuit that simultaneously services several users. The capacity of the circuit is shared among participants on the basis of time division, using a technique known as time-division multiple-access (TDMA), as shown in Figure 2.

FIGURE 2
TYPICAL PARTICIPATION IN THE JTIDS



Each participant in the JTIDS network is equipped with a synchronized clock, and is assigned a sufficient portion of the system capacity to accommodate the number of messages likely to be required by his mission. During his assigned transmit times, each user broadcasts data into a commonly accessible communications data stream, represented by the ring in the figure. All other elements can extract information of the type they require by continuously monitoring and sampling the data base. Digital processing provides each participant with selective access to all of the information generated by the other elements by applying fixed and variable filters to incoming messages.

Participants who have information will broadcast that information routinely into the net without needing to know who the recipients may be; tactical elements needing the data will extract it from the net without needing to know who furnished it. The user does not have to request information from a specific party, or wait until he is notified of information important to his mission; instead, he decides what category of data he wants - such as hostile aircraft within a 50-mile range - and he will receive everything the system has in that category.

1.5 JTIDS Messages

An individual message inserted into the data base may be intended for general dissemination or it may be addressed to one or more discrete recipients. Consequently, JTIDS simultaneously effects communication from one element to many elements, from one to one, from many to one, or from many to many.

JTIDS will support two basic message types - formatted and unformatted:

(1) Formatted messages are highly structured and conform to predefined JTIDS formats. Each information bit has specific meaning; therefore, a large amount of information can be compressed into a single time increment. Formatted digital messages are expected to replace much of the traffic now carried on UHF voice channels.

(2) Unformatted messages transmit data that does not fit the standard message formats, such as digital voice and teletype.

One ground station provides translational stability, but allows the grid to drift in rotation around that station. The mobile primary Users serve as the rotational reference for the community.

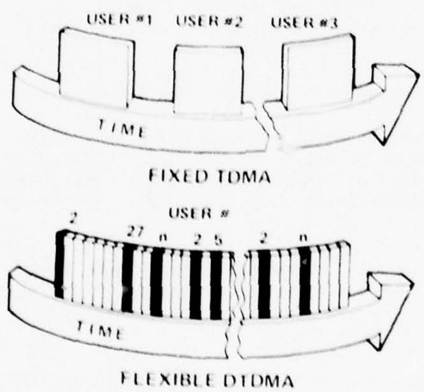
1.6 Development Program

The JTIDS development program is constrained to ensure that its hardware will be interoperable with existing and planned systems. JTIDS is being developed in two phases:

(1) Phase I will produce, in the near term, a comprehensive system for distributing information to support the command, control, and execution of tactical missions. It is using TDMA to create multiple subscriber nets; several such nets can be operated in the same geographical area through the use of code division. All elements transmitting on a single net must be accurately time-synchronized. Phase I terminals include TACAN and relative navigation.

(2) Phase II terminals will be compatible and interoperable with Phase I architecture and equipment. Phase II is a set of expanded capabilities to enhance the Phase I technology and will provide increased information transfer capacity and further consolidation of communication, relative navigation, TACAN and IFF functions. An alternative system architecture of distributed time-division multiple-access (DTDMA) is being evaluated for Phase II. DTDMA provides greater flexiblity than TDMA in the processing of various data channels and can more efficiently adjust to transmission timing and protocol requirements of existing tactical data links when required to emulate existing standards such as TADIL A. Transmissions can be at relatively slow rates and interleaved with the transmission of other units. The distribution and interleaving of pulses from a number of user transmissions will result in a lower probability of direction finding by the enemy on any one transmitter. This is illustrated in Figure 3.

FIGURE 3
JTIDS ACCESS TECHNIQUES



The JTIDS operational requirements which form the basis of these capabilities are summarized in the next section. The JTIDS ICNI capabilities have evolved from approximately 12 years of basic ICNI research and development by the U.S. Department of Defense.

2.0 JTIDS OPERATIONAL REQUIREMENTS

The JTIDS operational requirements address the 19 areas shown in Figure 4. This section summarizes the requirements and the corresponding JTIDS features which address the requirements:

FIGURE 4
JTIDS OPERATIONAL REQUIREMENTS

SYSTEM CAPACITY	GRACEFUL DEGRADATION
A/J - LPE - LPT	RANGE ACCURACY
MULTIPLE SIMULTANEOUS NETS	ICAO/NATO COMPATIBILITY
COST EFFECTIVE	CONNECTIVITY
JTIDS I/II INTEROPERABILITY	SURVIVABILITY
VARIABLE SYSTEM ARCHITECTURE	NAVIGATION
SECURITY	IDENTIFICATION
MODULAR CONSTRUCTION	MESSAGE STANDARDS
JOINT OPERATIONS	RECONSTRUCTABILITY
ACCESS	EVOLUTIONARY

(1) System Capacity - High volume information capacity is required to distribute a large volume of information throughout an entire tactical theatre. This is addressed through the employment of digital transmission and formatted data in JTIDS resulting in increases in capacity of at least 1000 to 1 over conventional analog voice transmission.

(2) A/J - LPE - LPT - Anti-jam, low probability of exploitation, low probability of targeting is an obvious need which is addressed through the employment of spread-spectrum modulation over an extremely wide bandwidth in JTIDS. Burst transmission pseudonoise and frequency hopping techniques are all used.

(3) Multiple Simultaneous Nets - Multiple simultaneous nets are required to connect different tactical forces which have different ICNI needs. JTIDS provides this capability through a combination of time division, frequency division and code division multiplexing.

(4) Cost Effectiveness - Cost effectiveness is an essential requirement of any new system. The combination of many functions, including TACAN and IFF in a single system provides a more cost effective implementation than implementing these functions in several "black boxes."

(5) JTIDS I/II Interoperability - JTIDS I/II interoperability is essential to avoid obsolescence of JTIDS terminals produced early in the development cycle. JTIDS Phase II terminals under development provide all of the Phase I capabilities simultaneously with the expanded Phase II capabilities.

(6) Variable System Architecture - System architecture flexibility is essential to meeting the changing needs of the tactical community. The TDMA and DTDMA architectures in JTIDS and the JTIDS ability to accommodate a variety of message formats and protocols provide this necessary flexibility.

(7) Security - A secure data unit is an integral part of the JTIDS design.

(8) Modular Construction - Modular construction is required so that JTIDS terminals can meet a wide variety of platform needs. Major modules of JTIDS terminals which can be changed or removed depending on platform requirements include high power amplifiers, data processors, and TACAN and IFF processing modules.

(9) Joint Operations - JTIDS is designed to be compatible with both NATO and U.S. joint service message standards incuding TACS/TADS, NADGE and UKADGE. Interfaces are under development for all of these and will be discussed in subsequent sections of this report.

(10) Access - Rapid access to the system is essential for real-time tactical data exchange. The chocie of TDMA and DTDMA architectures allow access times as short as 6.5 milliseconds and as long as 12.8 minutes depending on platform needs.

(11) Graceful Degradation - Graceful degradation of terminals and the system is essential to the survivability of the system. Examples of design for graceful degradation include the high power amplifier, which will only reduce power output in the event of transistor failures and the ability of the system to provide all of its ICNI functions, even if only two force elements remain in a net.

(12) Range Accuracy - JTIDS provides sufficient range accuracy between platforms to provide relative navigation accuracies to the level necessary to support tactical weapons delivery.

(13) ICAO/NATO Compatibility - ICAO/NATO compatibility are provided by JTIDS waveform and system architecture. The JTIDS waveform has undergone extensive testing and has been proved to operate compatibly in the 960 - 1215 MHz frequency band with existing TACAN, DME and IFF services in the band. JTIDS can provide interfaces and data exchange for NATO standard messages such as Link 1.

(14) Connectivity - The DTDMA and TDMA architectures provide connectivity levels between using platforms ranging from functional private subnets involving a few platforms up to total connectivity over a single TDMA net.

(15) Survivability - JTIDS is as survivable as the platforms themselves. It requires no control nodes or information choke points for the system to operate and provide all of its ICNI functions.

(16) Navigation and Identication - JTIDS navigation and identification capabilities and requirements were discussed previously.

(17) Message Standards - JTIDS must be compatible with existing joint service and NATO message standards for tactical data. Terminals and interfaces under development will enable the transmission of current messages such as TADIL A, B, C and Link 1 over the JTIDS network giving the data terminals which use these messages the link reliability, anti-jam and LPE/LPT features of JTIDS. In addition a common joint service message standard known as TADIL J is under development. The purpose of the TADIL J development is to provide a joint service message standard that takes full advantage of the JTIDS tactical data exchange capabilities and will provide interoperability among different message standards through Adaptable Surface Interface terminal (ASIT) currently under development.

(18) Reconstructability - JTIDS nets must be reconstructable even under the loss of the most critical user elements. This capability was discussed under survivability and graceful degradation.

(19) Evolutionary - The wide range of capabilities provided by the JTIDS terminals allows operational implementations starting with the provision of AJ/LPE/LPT and higher link reliability to existing data links evolving to a fully connected, interoperable ICNI system.

3.0 PHASE I, II COMPATIBILITY

In addition to the Operational Requirements listed in the previous section, there are specific requirements governing the two JTIDS program phases. These requirements are listed in Figure 5 and further elaboration on JTIDS terminal and system capabilities are given in the system description, hardware development and operational use sections of the AGARDograph. The last two Phase II requirements are worth elaboration in a separate section and they are covered here.

FIGURE 5
PHASE I/PHASE II REQUIREMENTS

PHASE I

(1) PROVIDE A BASIC DIGITAL JAM-RESISTANT TDMA COMMUNICATION SYSTEM AND A RELATIVE NAVIGATION CAPABILITY.

(2) ESTABLISH A SYSTEM WAVEFORM FOR PHASE I REQUIREMENTS WHICH ADDRESSES PHASE II.

(3) EMPLOY TDMA ARCHITECTURE WITH 25 TO 64 KBPS SINGLE NET DATA RATE.

(4) PROVIDE FULL INTEROPERABILITY WITH C&C SYSTEMS: NADGE, NTDS, TACS, TSQ-73, ETC.

PHASE II

(1) PROVIDE GROWTH IN DATA RATE, MULTIPLE NETTING, AND VERSATILITY TO INCORPORATE OTHER FUNCTIONS (IFF, GPS, DABS, ...).

(2) DEVELOP IN THREE PHASES: CONCEPT DEFINITION, CONCEPT VALIDATION, FULL SCALE DEVELOPMENT.

(3) BACKWARD COMPATIBLE WITH PHASE I.

(4) TDMA NOT REQUIRED IN ALL MODES

Backward compatibility with Phase I requires that Phase II developments be able to provide all of the Phase I TDMA modes of operation in addition to any new features which provide the growth in data rate, multiple netting and versatility to incorporate other ICNI functions. The expanded capability is provided by the addition of other terminal modes of operation which can operate simultaneously with the TDMA capability. This capability avoids the problem of obsolescence of Phase I equipment when Phase II is deployed.

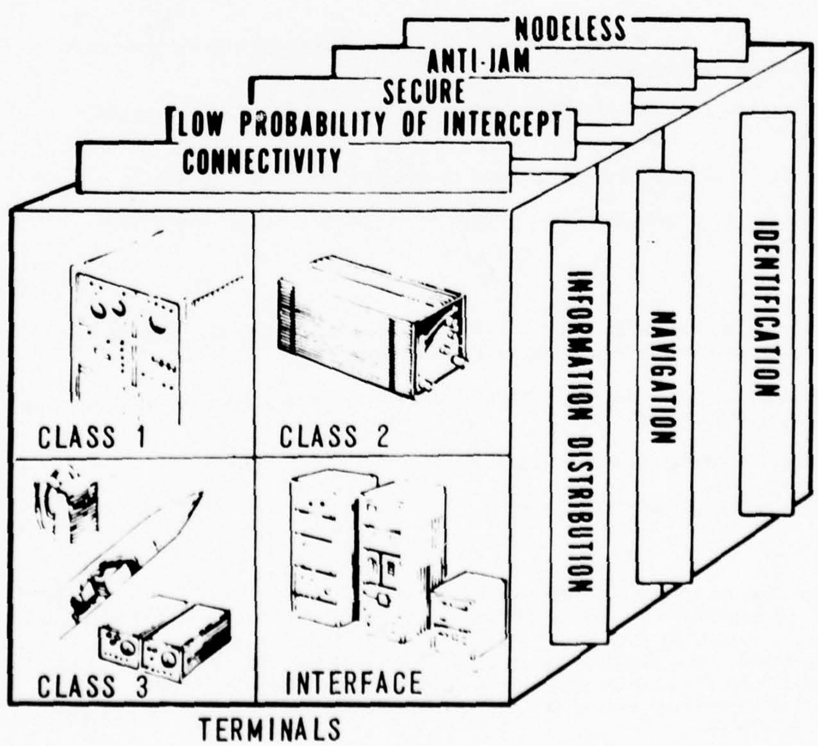
4.0 TERMINAL CLASSES

Three classes of terminals are under development in the JTIDS program, all three of which are applicable to both Phase I and Phase II of the program. Class 1 terminals or command terminals are intended for major air and surface command platforms such as the E-3A Airborne Warning and Control System, ground Tactical Air Control Centers and Navy Tactical Data System Command Ships. Class 1 terminals incorporate 1 kilowatt power amplifiers. Class 2 terminals or tactical terminals are intended for tactical aircraft, ships and other similar applications and incorporate 200 watt power amplifiers. Class 3 terminals have yet further limited capabilities and include several subclasses for applications such as manpacks, small ground vehicles and boats, and missile and RPV guidance. Developments in each of these classes are discussed in more detail in subsequent sections of this AGARDograph.

5.0 IMPLEMENTATION CONCEPTS

JTIDS is being developed to provide an integrated communications, navigation and identification capability to a wide variety of tactical platforms. Each platform equipped with JTIDS will be able to select the data it requires to perform its mission, and each platform will be able to provide the necessary data via JTIDS to support enhanced interoperability among other, intelligence, command and control and mission execution elements. JTIDS Class 1, Phase I terminals are currently being installed on the E-3A Airborne Warning and Control System Aircraft to provide a data link for exchange of position and track data and potentially provide digital voice, relative navigation and vectoring capabilities for control of tactical aircraft. Integration studies are underway and interfaces under development for providing the entire realm of JTIDS Class 2 terminal capabilities to the F-15, F-14 and E2C aircraft and NTDS equipped ships. ASIT's are under development to provide interoperability between the E-3A and US Air Force ground command centers. In short, JTIDS provides an order of magnitude increase in capabilities over current limited "black box" data links, navigation systems, and identification systems and provides the platform developer with a choice of implementations and interfaces from which he can select to fulfill his mission requirements. The system features, terminal classes and functions that the platform developer can choose from are summarized in Figure 6.

FIGURE 6
JTIDS SERVICE FOR TACTICAL OPERATIONS



NAVIGATION ARCHITECTURE
Edited by
LAWRENCE NEWMAN
COMMUNICATIONS AND NAVIGATION
TECHNOLOGY DIRECTORATE
NAVAL AIR DEVELOPMENT CENTER
WARMINSTER, PA. 18974
USA

SUMMARY

JTIDS provides precision range measurement, position data interchange and data processing functions which form the basis for its relative navigation capability. The JTIDS relative navigation system establishes the tactical grid, resolves the grid lock problem, improves on board system accuracy and shares community navigation resources. The current ADM JTIDS program is geared toward the demonstration of fundamental capabilities within the framework of a more extensive set of goals which reflect the full potential of relative navigation.

I. INTRODUCTION

Lawrence Newman, Naval Air Development Center

In essence, navigation is a process which stems from an ability to assign coordinates in a definable grid system to platforms or other objects of interest. JTIDS, the Joint Tactical Information Distribution System, has this inherent capability by virtue of its precision range measurement, position data interchange and data processing features.

JTIDS also disseminates its navigation data to the entire tactical community, thereby providing an essential consistency of position location to each of its elements. This consistency, as stated in (1), applies not only to the positions of each member but to all data derived in the community from on-board sensors and fleet tactical data interchange systems. This permits the acquisition of precise fire control solutions based on sensor data derived from multiple platforms and stations, i.e., data "fusion".

***

Our discussion begins with an exploration of the objectives of the JTIDS navigation function and a description of the basic principles involved. This will establish a frame of reference in which to hold the remaining segments which explore the functions of the current ADM (Advanced Development Model) system and the technological issues of JTIDS architecture which mold the community's basic operational potential. The paper closes with a brief look at the impact of this new capability as seen through its many potential applications.

Our basic purpose, therefore, is threefold: (1) To portray the basic features, functions and utility of this new capability; (2) To convey the complex architectural design issues whose resolution plays a vital role in defining the system's capacities and limitations and (3) To display the current system design juxtaposed against the full potential of JTIDS navigation. This should aid in the perception of both the program's achievements and the goals it wishes to attain.

II. OBJECTIVES OF RELATIVE NAVIGATION

James D. Bivin, Naval Air Development Center

It is said that knowledge is power. Certainly, in tactical operations, the commander with best knowledge of the disposition of friendly and enemy forces has a decided advantage. For this reason, the introduction of tactical data links has effectively increased the fire power of tactical vehicles through a more effective utilization of available intelligence and reliable real-time semi-automated command and control procedures.

These improved capabilities have also introduced a special technical challenge. Navigation errors cause discrepancies in reported position, resulting in the same target being reported in more than one place when held by more than one sensor. This produces the "grid lock" or "track correlation" problem with which engineers have struggled to cope for the past decade.

This is precisely the problem that JTIDS relative navigation was designed to overcome. The new TDMA (Time Division Multiple Access) spread spectrum data links have an inherent range measurement capability which can be used recursively to eliminate position discrepancies between participants in the relative navigation community and then to reduce the actual sources of on board navigation system position, velocity and heading errors.

JTIDS relative navigation is designed to operate with a minimal dependency on sources of accurate absolute geographic position. However, once the participants in the relative navigation community have accurately established their relative position (i.e., their displacements from one another), they can share their geographic navigation resources and thus improve the absolute position accuracy of members equipped with less precise geographic position references. In essence, the relative navigation grid can be used to relay the geographic position of the better equipped members.

$$\begin{bmatrix} \sin\Theta_4\cos\psi_4 & \sin\Theta_4\sin\psi_4 & \cos\Theta_4 & -1 \end{bmatrix}$$

The JTIDS relative navigation technique, as professed in the Hippocratic principle that the treatment shall not leave the patient worse, takes care not to degrade the performance of any members of the community. Recognizing that the tactical grid position may only be loosely related to geographic position, the relative corrections are carried as offsets or corrections to be applied to the on-board navigation data in a feed forward mechanization (i.e., corrections are applied to the navigation system outputs as opposed to feeding back corrections to reduce system error sources). The option of applying these corrections to the on-board geographic navigation system in a feedback mechanization may be exercised if, for example, an inertial system is known to be poorly aligned.

Not only does JTIDS relative navigation provide a large payoff at little cost, but it also compliments the new GPS (Global Positioning System) precision geographic positioning capability as though two interlocking pieces of a jigsaw puzzle. GPS represents a tremendous advance in the practice of navigation, using orbiting satellites to provide pinpoint accuracy anywhere in the world using fairly simple receivers. However, due to the extreme distance between the satellite and the receiver, only a relatively weak signal is available and occasional loss of signal is expected. In this situation JTIDS relative navigation, with its higher transmitted power and shorter distance to the transmitter, makes an excellent relay. Furthermore, it is possible to give the benefit of GPS to all participants while only equipping selected high value platforms with GPS receivers.

The JTIDS relative navigation capability is therefore rather elegant in its simplicity. Inherent within the data link, it makes use of range measurements which are the byproduct of data transmissions to solve the "grid lock" problem. Since its range measurement capability is implicit, the incremental cost of providing relative navigation is minimal. In establishing the tactical grid, resolving the grid lock problem, improving on board navigation system accuracy, and sharing community geographic positioning resources, we have defined the basic objectives of JTIDS relative navigation.

III. BASIC PRINCIPLES

Lawrence Newman, Naval Air Development Center

By virtue of its capacity to

- determine precise range between members of the community
- transmit position data among participants
- digitally process information

the JTIDS relative navigation capability provides the joint services with a cohesive basis for the formation of a community navigation net through the synergistic combination of multiple platform navigation resources and allied data processing capabilities.

The current phase of the JTIDS development program for relative navigation is geared to the establishment and demonstration of fundamental capabilities within the framework of a more extensive set of goals which reflect the full potential of relative navigation. This prudent approach is consistent with the scope and objectives of an ADM program in that its aim is to provide the necessary level of confidence in basic concepts which will permit the development program to advance to the EDM (Engineering Development Model) phase.

As such, the purpose of this segment of the paper is to provide a larger frame of reference in which to hold the remaining segments of this paper which are mainly devoted to descriptions of the current ADM system. This should provide sufficient clarity to understand where the program is now and to see the full potential of the relative navigation capability.

***

In its largest sense, the JTIDS community navigation net is comprised of navigation references and a navigation processor. The navigation references may consist of the JTIDS antenna/receiver (a range measurement Navaid) plus navigation equipment extant on member and cooperating platforms - e.g., dead reckoning systems (inertial, doppler, etc.), geographic and relative position references (GPS, radar, etc.), altimetry and other special references.

The navigation processor integrates all or a subset of the measurements derived from the navigation references to provide the platform relative navigation capability. In addition to the navigation measurements, the relative navigation algorithms employ the following inputs:

- measures of navigation quality received from participating JTIDS platforms in the community which are used to select and/or weight received navigation measurements.
- on-board operator commands which establish the role of each member in the community's relative navigation architecture as well as providing prerequisite system initialization, parameter insertion and other functions as required.
- position coordinates of community members (source: JTIDS messages) and non-community members (source: on board sensors, community members' sensors and cooperating community member sensors) for insertion into the grid.

***

The integrated relative navigation system employs these inputs to perform five basic functions which may be termed

- GRID SYNCHRONIZATION
- GRID ACQUISITION
- GRID UPDATE
- GRID EXTENSION
- GRID INSERTION

The paragraphs below will define these functions. (Refer to Figure 1.)

Grid Synchronization is the establishment of a time-of-arrival (TOA) measurement capability from which accurate range measurements may be derived. Grid Synchronization is established by a Synchronization Routine which may, in general, utilize "active" (two-way) techniques such as round trip timing (RTT) or "passive" techniques such as pseudo-ranging (use of known position to derive clock corrections; see discussion of Grid Acquisition) or the use of independent time references.

In a typical community it is anticipated that, initially, a small group of members will engage in synchronization using active techniques. Thereafter, additional members of the community will synchronize using passive techniques to minimize transmission load. The availability of independent time measures will, of course, supersede these considerations.

Grid Acquisition is the establishment of the on-board navigation grid when the system is energized or upon entry into a community. Grid Update is the maintenance of on-board grid alignment during mission operations. Grid Acquisition and Update are accomplished by applying a subset of the navigation system measurements provided as inputs and selected via a Source Selection Routine to establish and maintain convergence of the Navigation Algorithm so as to provide Corrections and Calibrations to grid coordinates and to the navigation system measurements.

The Source Selection Routine prefilters the multiple measurements available to a platform in accordance with indicators of measurement quality received from these sources. If sources are expected to be less frequent, the Source Selection Routine can be simplified to selected weighting factors such as that used in Kalman filter algorithms.

The Navigation Algorithm uses the selected measurements to estimate corrections to relative and absolute grid coordinates, time of arrival measurements ("passive" clock synchronization) and, for high precision applications, errors in measurements from the selected sources (e.g., gyro biases) and estimates of various navigation-related environmental parameters (e.g., wind speed, ocean current). It is anticipated that a Kalman filter recursive algorithm such as that employed on modern integrated navigation sysetms will be used for this function in most applications.

The Corrections and Calibrations Routine provides the iterative application of Navigation Filter estimates of platform grid coordinates and (when available) geodetic coordinate errors to the navigation processor's stored values of these quantities. The application of time-of-arrival calibrations, measurement source corrections and navigation-related environmental parameters is accomplished as required by other routines. The Corrections and Calibrations Routine shall also retain and update parameters relating platform, grid and geodetic coordinates.

Grid Extension is the establishment of grid coordinate relationships between the on-board navigation grid system and that of members of an independent community. Grid Extension is accomplished by an Intergrid Maintenance Routine.

In general, the Intergrid Maintenance Routine shall establish the coordinate relationships between both nodal and non-nodal communities. A separate intergrid parameter estimation filter establishes the relationships based on measurements between members of each community. Upon convergence, these estimates are applied in a grid coordinate conversion algorithm which homogenizes inputs received from outside communities so that they become indistinguishable from measurements derived from similar references within the community. The architectural design for each community (nodal vs. non-nodal, primary community references for grid definition, geodetic position, time, etc.) is a subject which is treated in a subsequent segment of this paper.

Grid Insertion is the assignation of grid coordinates to community members and non-community members to provide the following capabilities:

- precision grid position located for force disposition monitoring and mission coordination
- correlation of target and weapon coordinates
- precision aircraft vectoring
- over-the-horizon targeting
- geodetic position reporting based on surveyed geodetic station sites or on-board Navaids.

Grid Insertion is accomplished using a Locator Routine which uses coordinates received from community, extra-community and independent cooperative platforms.

Thus $H^{\wedge}Q^{-1}H$ is singular

Q.E.D.

The Locator Routine updates stored coordinates using geographic and grid position information, sensor detected range, bearing, mark-on-top, etc., course and speed estimates from tracking algorithms and other sources. Grid positions thus assigned are transmitted to other CNI routines as required.

***

The on-going program, by comparison, has thus far concentrated on the development of a prototype JTIDS-integrated relative navigation capability which employs JTIDS range measurements to bound inertial system error propagation using a Kalman filter algorithm. Expressed in terms of the preceding material, the ADM system provides the following:

Measurements:

(1) JTIDS range
(2) Inertial system position, velocity, heading
(3) Ground Station position

Other Inputs:

(1) Measures of position, time and azimuth quality (conveyed via JTIDS messages)
(2) Operator-inserted commands to establish role in the community as either a controller (of grid coordinates, geodetic coordinates, and/or time references) or as a user
(3) Initialization, parameter insertion and other specialized activation functions

Position Coordinates:

(1) Grid coordinates of community members (conveyed via JTIDS messages)

Grid Synchronization:

(1) "Active" RTT
(2) "Passive" pseudo-ranging using Kalman filter

Grid Acquisition and Update:

(1) Source Selection Routine identifies subset of received range measurements using algorithms based on a comparison of the quality levels of received data and on-board navigation information.
(2) Kalman Filter Navigation Algorithm which uses range measurements and ground station position (when available) to estimate corrections to relative and absolute grid coordinates, time-of-arrival measurements ("passive" clock synchronization) and inertial system parameters.

(3) Corrections and Calibrations Routine consisting of feed-forward corrections to the quantities cited in (2).

The Grid Extension and Grid Insertion functions as well as other subfunctions identified in this paper will be addressed in the EDM phase of the program wherein the assignment of receivers to specific platforms for designated missions to achieve prerequisite accuracies will provide the basis for the designation of the following constraints to the design problem:

operating environment

navigation measurement sources

quantity of community and non-community elements

measurements available from or to the above elements

necessity for grid extension

data processing resources

method of integration with existing grid insertion capabilities (NTDS, ATDS, MTDS, etc.)

These shall provide the basis for the formulation of specific operational software packages and the establishment of the JTIDS community navigation net architecture.

***

In the sections which follow, we will endeavor to explore in greater detail the Grid Synchronization techniques (Section IV), relative navigation architectural issues including community protocol (Section V) and community structure (Section VI), and data processing functions (Section VII) as they apply to the current system.

IV. SYNCHRONIZATION AND RANGING

Robert Stow, Singer-Kearfott Corporation

The Grid Synchronization function establishes the time of arrival (TOA) measurement capability from which accurate range measurements may be derived. This segment of the paper provides a detailed description of the techniques designed to perform this function in the ADM JTIDS terminal.

***

The JTIDS terminals require accurate time synchronization to permit receipt and transmission of messages within assigned time slots. Time synchronization is also required to determine the time of arrival (TOA) of received messages. TOA provides a measure of range between community members, when scaled by the speed of light, for use in relative navigation. The synchronization techniques utilized are applicable to large community sizes operating within a 300 nautical mile line-of-sight. Both active and passive synchronization modes are utilized to meet the requirements of user community members in both a radiating and non-radiating environment. Passive synchronization, in particular, can be used to conserve the number of slot assignments required to support the sync function. In addition, each user is capable of switching from active to passive sync modes and adapting to variable synchronization update rates from selected community members in passive or active modes.

As shown in Figure 2 user time synchronization is a two step process: (1) initial net entry and coarse synchronization followed by (2) fine synchronization. Figure 2 shows coarse synchronization being achieved by a synchronizing member with the receipt of a net entry message transmitted from a synchornized member in slot N. Following this coarse time slot sync, Round Trip Timing (RTT) Interrogation Messages may be sent by the user and the responses used to determine a measurement of slot time synchronization error as shown in slot N + M. The measurement of time slot error may be computed by taking one half the difference between the RTT interrogation TOA measured by the synchronized member with respect to the beginning of slot N + M, and the RTT replay TOA measured by the unsynchronized member with respect to the mid slot reference time. The RTT interrogation TOA is transmitted to the unsynchronized member in the RTT reply message. By repeatedly correcting the user's clock time reference with these time error measurements, fine synchronization is achieved as shown in slot N + K where now the normal transmitted message TOA measurements are directly proportional to range between the two community members. This fine synchronization procedure is used in active synchronization modes. In passive modes the RTT messages are not utilized and fine sync is achieved by utilizing normal transmitted messages from other synchronized community members.

A hierarchy is established within the JTIDS community for message synchronization. One member is arbitrarily defined as the Time Reference, and this member makes no corrections to his clock while operating within the community. This member serves as the time standard. This unit transmits net entry messages in preassigned time slots and automatically transmits the highest time quality level via the position and status messages (P-message). Users entering the net will obtain initial net entry, coarse synchronization and fine synchronization from this relative time reference and will adjust their time quality to be one less than their synchronization source. As a community grows, entering terminals can obtain synchronization from any user already synchronized within the community and within line-of-sight of the entering unit. The time quality hierarchy established and maintained within the community automatically allows for users to maintain synchronization to the best sources within line-of-sight of each respective unit.

The initial net entry process is employed by entering terminals to automatically achieve synchronization with community time to a degree sufficient to receive messages. When operating in the JTIDS anti-jam, secure modes, wherein the message coding and/or RF frequency pattern change as a function of slot time of day, initial net entry is performed by seeking to detect a transmission from a unit within the community that is transmitting in net entry slots. The entering unit initiates reception enough in advance to detect a specific transmission that will occur in the future. This look-ahead time is based upon the entrant's uncertainty in initial estimate of system slot time. All user terminals initialize their slot counters to their best estimate of GMT (Greenwich Mean Time) and indicate the maximum uncertainty this estimate may have. The entering unit listens exclusively for a specific net entry transmission for an amount of time equal to twice this time uncertainty. If at the end of that period, an error free message has not been detected, the process is repeated. If, when attempting net entry, a message is detected which is subsequently determined to be in error, the terminal will conclude the current attempt and repeat the process. Once a radio input message is correctly received, system time will be known at the receiving terminal with an uncertainty equal to the propagation time of the signal between transmitting and receiving terminals plus the community time sync error of the transmitter. When operating in JTIDS non-secure modes, an initial knowledge of system time is not required for message reception, since these modes utilize fixed modulation codes.

Following initial reception of an error free message, the terminal declares itself to be in coarse synchronization and automatically starts the fine synchronization process. There are three operator selectable fine synchronization procedures, called Passive I, Passive II, and (active) Round Trip Timing (RTT) sync. The first two procedures require no transmission by the synchronizing unit. Only transmission from source community members are needed. The third procedure requires interrogation transmissions to be made by the synchronizing unit to another addressed terminal. The addressed terminal responds with an RTT reply message which contains the time-of-arrival of the RTT interrogation message. The interrogating terminal can compute from the RTT reply message a correction to its estimate of system time as discussed above.

service in addition to the navigation service. These studies can ... sidered to be the genesis of the integrated communication-navigation-identification (ICNI) functions later embodied in JTIDS.

The RTT sync mode makes use of a digital filtering technique for correcting relative clock and oscillator errors within the TDMA transceiver. The technique employs a linear filter with time varying gains to estimate relative clock and oscillator frequency errors from measurements of community time relative to the RTT interrogated unit. This iterative clock estimation procedure allows rapid time synchronization to accuracies which will allow accurate interpretation of received message TOA as range for relative navigation. Also, normal JTIDS message transmissions can now be made by the synchronizing terminal and received by other synchronized community members without overlapping slot boundaries. The determination of which community members to interrogate when in this synchronization mode relies upon the identification of sync reference sources, based on received P-messages whose reported time qualities are better than the synchronizing unit's own time quality. During periods where no sources are available, the linear filter algorithm degrades its own unit time quality estimate consistent with its clock drift model. A maximum rate is placed upon the transmission of RTT interrogations so as to limit user transmit slot utilization for this function.

The Passive I sync mode makes use of received P-message data and associated message TOA to compute user clock error. This estimate of user clock error is based upon the received P-message position report data, message TOA, and user position supplied external to the terminal. As in the case of RTT sync, the system time measurements are used to update a linear filter model which estimates user clock bias and oscillator frequency errors relative to those units whose P-messages are processed. The determination of which source P-message to use for update is based upon identifying those sources of better time and position quality than the synchronizing terminal. Unlike the RTT sync mode, which is independent of any community knowledge of user positions, the Passive I sync mode requires use of both the position of the source and the position of the synchronizing unit. The accuracy of the clock bias error estimate is therefore primarily limited by the relative position errors of the users.

The Passive II fine sync mode utilizes received P-message position reports and message TOA data from selected sources. This TOA data consists of the ranges between the passive sync user and each transmitting unit, plus the user's clock time bias with respect to system community time. This data is mixed with vehicle dead reckoning subsystem data in the terminal relative navigation filter algorithm to determine both user clock errors and estimates of user position, velocity, and dead reckoning subsystem attitude errors. The dead reckoning subsystem provides an extrapolation or "flywheel" capability between TOA measurements in the position domain similar to the time extrapolation capability provided by the oscillator. Using this fine sync mode accuracies similar to the RTT sync mode are achievable since position errors are also calibrated simulataneously with clock errors.

As discussed in (2), the ultimate capability of the relative navigation function is limited by the TOA error and the computational error of the processor. Although GDOP (geometric dilution of precision) is an additional consideration, the dynamics of platform motion will reduce GDOP in ordinary situations as a result of the Passive II filtering process. The TOA error, in turn, is ultimately limited by the short term stability of the terminal clock, propagation anomalies in the atmosphere (actually, the residuals of propagation corrections; see (3) for a more complete discussion) and multipath effects.

V. POSITION REPORTING PROTOCOL

E. A. Westbrook, MITRE Corporation

The architectural design of the current JTIDS relative navigation community is based upon a system of platform protocols wherein members of the community are assigned specific roles which determine who updates whom and which platforms are to serve as primary references for relative positon, geodetic position and time. This segment of the paper provides a brief discussion of protocol mechanics as well as a definitive description of the coordinate frames in which the system operates, the reporting rates employed by the message structure, and the manner in which members of the community acquire the grid.

***

A. Protocol

Each position message contains four quality fields: time quality, geodetic position quality, relative position quality and relative azimuth quality. These are transmitted to support the hierarchial community structure by serving as badges of rank. They are derived from and are translatable into time and position variances and represent, therefore, the estimated uncertainty of the reported positions and clock synchronization accuracy. They are used by the receiving unit to select the reports to be processed for relative navigation filter updates and to establish the observation variance in those filter update computations.

The source selection protocol is intended to serve two basic purposes: maintenance of system stability and optimization of accuracies within the limited processing time available. This is done by establishing the general rule that only sources of higher quality (lower variance) may be used for filter update. (Certain classes of user are allowed to use others of equal quality.) This prevents reciprocal or circular updating (feedback). The second objective is achieved by selecting from the set of sources of higher quality the best subset that can be processed within the time available. Typically this is three or four observations each 10 to 15 seconds in the present JTIDS units.

At the top of the hierarchy stand Position References, Navigation Controller and Time Reference. The Time Reference is assigned a time quality of 15. The Navigation Controller is assigned a relative position quality of 15. Units having known geodetic position to within 50 feet are designated Position References and assigned a geodetic position quality of 15. Other units determine system time (clock synchronization) and position by Kalman filter combination of inter-element ranges with dead reckoner outputs and establish their own quality levels from the filter covariance matrix, always restricting the quality transmitted in the P-messages to a value not greater than that of the source(s) to preserve the hierarchy.

* An exception being the most advanced ICNI multifunction architectural development in the other-than-L-band portions of the spectrum. This part of the program is still peculiarly Navy in its formative stage, and is known as the Tactical Information Exchange System (TIES), see Section III. A.

Position References, Navigation Controller and certain other select units called Primaries use round trip timing for clock synchronization. Other units synchronize passively using multiple pseudo-range measurements.

The various classifications of units refer only to operating modes and do not denote differences in user equipment. Any unit may act in any capacity.

***

B. Reporting Rates

There is no fixed reporting rate in the JTIDS. It is individually set in each unit as a function of both the tactical mission of the unit and the role of the unit in the relative navigation community. A typical position reporting rate is once each twelve (12) seconds. This number is roughly comparable to typical radar position reports (scan time) and to the time required to process a set of three sources required for an unambiguous passive determination of position and time in the navigation filter.

***

C. Coordinate Frames

Two coordinate frames are reported in the P-messages (see Figure 3). The first, in the main body of the message, is the familiar geodetic latitude and longitude system. The basic spheroid underlying is defined to have an equatorial radius of 6,378,135 meters and a polar flattening of 1/298.26 as in WGS-72. This is the same basic spheroid adopted by the GPS (Global Positioning System). Latitude and longitude are reported in units of 1/512 minute of arc.

A second local, planar relative coordinate frame is reported in the P-message. The plane is defined as tangent to the earth's surface at its origin with the u-coordinate axis position east and the v-coordinate axis positive north. 'u' and 'v' are reported in units of 1/512 data mile. (A data mile is defined to be 6000 feet exactly, a unit commonly used in radar systems.) Grid north at the origin may be defined to lie at an angle "beta" to the local meridian. Beta is reported in units of 0.0125 milliradian.

The origin of the relative grid is initially defined by the Navigation Controller, but is not explicitly transmitted: the combination of latitude, longitude, altitude, u-position, v-position, and beta angle in each unit's P-message serves as implicit origin definition to other units.

***

D. *The Grid Acquisition Process

Relative grid navigation is initiated by one member, designated the Navigation Controller, establishing a reference point, or grid origin, at some arbitrary location. For convenience this point can be the location of a fixed ground site, a point on his trajectory, or any other point he designates. The controller then estimates his position relative to the reference point he has defined.

The Controller then navigates in this coordinate frame based on the velocity information provided by his on-board dead-reckoner equipment. The Controller also broadcasts his computed grid position as well as his geographic position via the data link. A prospective User, also equipped with the RF data link, can receive the Controller's reported positions (and/or the reported positions of other Users already operating in the grid) and with the DME capability of the data link also obtain a measurement of range to the Nav Controller (and/or other Users).

For initialization, the User computes an estimate of the grid origin from a combination of the reported geographic and grid positions received. Then using his own estimate of geographic position and the derived grid origin, the new User estimates his own grid position. These computations are performed only once, at acquisition. Following acquisition, changes in grid position are computed by integrating velocity from his on-board dead-reckoner.

Having the reported grid position of the Controller or other established community member and an estimate of his own grid position, the User can compute a predicted range to that other member. The actual measured range compared to the predicted range provides information as to his position error in the grid with respect to the observed member. By repetitively performing this operation, each time making corrections to improve his position estimates, the User achieves grid lock with the community. That is, he has now positioned himself in the established grid and possesses the capability to accurately navigate relative to the community between times when he receives range measurements.

The conventional trilateration requirement of ranges to at least two other members does not exist in this positioning algorithm. Although ranging to multiple members is advantageous, the on-board dead-reckoner provides sufficient additional information to correlate a number of range measurements to the same member. By virtue of the relative motion between the members, the ranges to a single member appear in different directions over a period of time. The integral of the dead-reckoner velocity links the range vectors together for uniquely positioning the aircraft in the grid.

* Special section contributed by W. Steele, Singer-Kearfott Corporation

Having acquired the relative grid, continual range measurements are used to maintain the grid lock of the User. In addition, the filter is designed to process three other observation types. Each of these is related to the use of geographic information for improved geodetic navigation. They are:

- Geodetic Fixes - Used when current latitude and longitude of the vehicle is available from some on-board source. This may be generated by anything from a "Fly-Over" of a known geometric location to the semi-continuous geodetic reference of a GPS receiver. This observation type is potentially the most beneficial for enhancement of geographic navigation.

- Offset Geodetic - Used when the position of any arbitrary point in the relative grid is known in geographic coordinates. The accurate relative grid data is used to translate the reference point to the self aircraft. This is a very powerful observation type since it allows the entire community to make use of geodetic data that was available to one of its members. This data could otherwise not be distributed with such precision.

- TOA Geodetic - Used when the geographic position of a TDMA equipped non-community member is accurately known. This process is similar to the relative grid range observations processing, but with geographic positions used in place of the relative grid positions for predicting range. This provides a geographic update capability to any TDMA user or to the community, even if the TDMA user is not navigating in the relative grid.

The filter observation selection logic automatically chooses among the four possible observation types based on availability, relative quality, geometry, and benefit to be derived. This observation logic assures stable, accurate, and non-fragmentary community operation for the large community as well as the limited one-on-one case.

VI. STRUCTURE OF THE NAVIGATION COMMUNITY

Melvyn S. Greenberg and Dr. H. James Rome, Dynamics Research Corp.

Section V described the manner in which the relative navigation community structure is regulated via a strict hierarchy of assignable and interchangeable roles. These cannot be specialized grid functions and operations which have been incorporated in order to sustain the accuracy and stability of grid navigation. The grid thereby provides a common datum for the exchange and synchronization of navigation information. It is the means by which the community participants may process relative measurements and lock their relative positions, relative velocities, and relative headings without the need for precise geodetic position.

This segment of the paper expands upon the mechanics of protocol to present the underlying architectural philosophy of the relative navigation community. Topics include grid hierarchy, system elements, grid stability, mixed communities, geodetic updates, interplatform error propagation, the usage of relative and geographic grids, and alternative organizational concepts. These are the architectural issues which bridge the gap between relative navigation concepts and mission capabilities.

***

A. Grid Hierarchy

On top of the grid hierarchy is the Navigation Controller who is responsible for establishing the location and orientation of the grid. The Time Base Controller is another grid role which serves in an analogous fashion by providing information in the time domain for synchronization of each member's time base.

The rules of grid membership have been developed in accordance with community stability, accuracy, and connectivity considerations. User membership is allowed for terminals participating as Primary or Secondary Users. Primary Users are a distinct class of users typified by an RTT active clock synchronization technique. Secondary Users, on the other hand, utilize passive synchronization techniques. As TOA measurements are widely used in positioning in the grid and geodetic frames, there is a strong dependence which develops between positioning accuracy and time synchronization accuracy. Consequently, in the hierarchy of position accuracy, the Primary Users are permitted in the Rel Nav protocol to assume the second highest status (after the nav controller) while Secondary Users are relegated to the next lowest level.

Each user, regardless of designation, determines the accuracy of its time synchronization grid position, geodetic position, and relative azimuth. The reported quality of this information is then used by the community as the basis for discriminating among potential candidate sources, for eventual selection and processing by the Rel Nav filter.

Another type of component within the grid is a Position Reference, whose role is to provide accurate geodetic position to other members of the community.

B. System Elements

There are three essential elements which contribute to a relative navigation system. One has already been mentioned; the grid frame. Another item is precision inter-unit range measurements, and the third element is a dead-reckoning navigation unit.

(3) Emulation of the TACAN interrogator function in all JTIDS tactical aircraft terminals to provide backward interoperability with current TACAN beacons and elimination of redundant TACAN "black boxes" on JTIDS equipped tactical aircraft.

- The concept of a grid frame frees the relative navigation problem from an absolute frame of reference.
- Precise inter-unit range measurements are the basis for accurate relative positioning schemes.
- A dead-reckoning navigation unit provides the short-term velocity and attitude stability for navigating between infrequent position updates.

Through the periodic filtering of range measurements and dead-reckoning data, these two independent navigation data sources are combined by the Rel Nav algorithms to produce accurate relative position, velocity, and heading. A significant consequence of Rel Nav performance is in the ability of one unit to point to another unit. This is the ability to precisely specify the orientation of a platform measured vector, such as a radar contact, to another remotely located Rel Nav unit. This result has been shown to be of fundamental importance to the accuracy of data exchange, target data hand-off, and general mission effectiveness.

***

C. Grid Stability

The characteristics of a community grid are determined in part by the two types of possible Navigation Controllers: ground stations or mobile units. Two ground stations or only a single mobile unit can be used to define a grid.

Mobile Naviation Controllers can define the grid location in two ways: at their local position or at some remotely determined offset point. A User enters or acquires the grid in only one way. A User must first extract the location of the grid origin which is implicitly embedded in the P-message report. The User then initializes its position in the grid on the basis of its imprecisely known latitude and longitude and the coordinates of the grid origin. Stationary Navigation Controllers provide a highly stable set of reference points which permit accurate and stable grid performance to be achieved, although geodetic position uncertainties may be large.

Mobile Navigation Controllers provide a suitable but less stable grid frame. The mobile navigation controller reports grid position on the basis of its dead-reckoning navigation system. No corrections are made to its reported grid position. Corrections to its reported geodetic position are allowed. Conditions of Nav Controller mobility and imperfect dead-reckoning measurements produce two distinct effects on community grid performance.

In any given period of time the distance covered by the mobile Navigation Controller can be likened to a fictitious baseline between two equivalent fixed Controllers. Since this distance is typically much shorter than ground station separations, short-term positioning accuracy relative to the Controller is adversely affected, and the accuracy is understandably very dependent on actual Controller motion.

A second somewhat more subtle mobile Nav Controller effect is introduced by its own dead-reckoning errors. When the Nav Controller has low frequency or bias like velocity errors, the reported position of the Nav Controller drifts away from its actual position. Users attempting to precisely navigate relative to the controller can only do so when the relative User-to-Controller velocity errors are zero. A User dropping out of LOS or a Nav Controller going radio silent will then not corrupt relative navigation accuracy. Long term steady state grid lock, for any user, under these conditions necessitates close-tracking of the characteristics of the Nav Controller's velocity errors.

*Although not necessary, it is beneficial for fixed ground stations to be included in the relative grid community. (Their geographic location need not be accurately known.) When observed by the mobile community members, these fixed stations serve to stabilize the grid relative to the earth. Two or more ground stations provide both translational and rotational stability. One ground station provides translational stability, but the community relies on other mobile members for its rotational stability.

Since the Navigation Controller takes no grid observations, ground stations are most effective when they are assigned the Nav Controller role. Then all mobile members will process the range observations to them. Once two (or more) ground stations are positioned in the relative grid, the baseline between them provides both a translational and rotational reference for the grid UV axes. Any community member in line-of-sight of the stations can process the TDMA system measured range in the Kalman filter and easily estimate his position in this fixed coordinate frame, or grid. Continued processing of observations, comparing predicted changes in position with observed changes in position, permits estimation of velocity in the UV grid.

During maneuvers, when acceleration is present, the angle between the inertial systems XY accelerometer axes and the grid UV axes can be quickly estimated by the filter in a manner conceptually similar to conventional acceleration matching. Ideally speaking, (assuming no inertial system errors), this now permits accurate relative position extrapolation using only the dead-reckoner inputs. Recognizing, however, that there will be inertial system errors, the continued observations allow for the recovery of these errors in much the same way as if there were a Doppler radar on board. The accurate grid position reference permits accelerometer tilt recoveries and, through gyrocompassing, allows geographic heading recovery. Thus much improved geographic navigation is also achieved.

* Special subsection contributed by W. Steele, Singer-Kearfott

Formatted messages are highly structured and conform to predefined JTIDS formats. Each information bit has specific meaning; therefore, a large amount of information can be compressed into a single time increment. Formatted digital messages are expected to replace much of the traffic now carried on UHF voice channels.

(2) Unformatted messages transmit data that does not fit the standard message formats, such as digital voice and teletype.

One ground station provides translational stability, but allows the grid to drift in rotation around that station. The mobile primary Users serve as the rotational reference for the community.

Consider a community composed of one ground station, the nav controller, and one mobile member. Nominally fixed to the earth, the grid will now appear to rotate (with respect to the earth) at a rate consistent with the navigation error of the mobile member. If there is adequate relative motion between the mobile User and the ground station, such as would be the case if the User flew a circle around the station, the drift can be minimized. By flying around the station it appears in a different direction each time a new observation is taken. This presents to the filter a "multiplicity of ground stations" allowing it to calibrate its navigation errors as before and thus reduce the rotational drift.

When more than one mobile member is present, the User with the best dead-reckoner and most relative motion with respect to the ground station will ultimately attain the highest quality level and serve as the rotational reference for the other members. The availability of geographic observations will also affect the convergence of the grid structure.

The third possibility, that of having no ground stations, allows the grid to drift freely in translation as well as rotation with the mobile navigation controller's error profile. It is, therefore, advantageous to assign as the navigation controller that community member with the most accurate navigation system. If the navigation controller is equipped with a very accurate geographic position reference sensor (e.g. GPS), the result is similar to having fixed ground stations in the community. The geographic observations allow the controller's filter to calibrate its navigation errors and hence limit the grid drift.

This phenomenon of a drifting grid does not directly imply degraded performance. Accurate relative navigation is still maintained among community members. Only the knowledge of the grid origin position relative to the earth degrades with time.

D. Mixed Communities

For homogeneous navigation communities, such as an all inertially equipped community, the ability for a User to remain accurately locked to a Controller is aided by the similarity of error signatures among grid elements. However, in Rel Nav communities with mixed types of dead-reckoning units (for example, some elements with Doppler systems and others with inertial systems), the maintenance of long-term accurate relative navigation does not occur automatically. It can only be achieved by the introduction of special mixed community software modifications.

Recognition of the constraints imposed by optimal relative navigation considerations in a mixed community give rise to some basic conditions which must be satisfied. Precise long-term navigation can only be assured by the representation in the Rel Nav filter of all significant sources of relative navigation error.

In the instance of an all inertial community it turns out that there are no additional error terms required to be introduced by the optimal solution. All terms are already included in the nominal filter formulations to satisfy the relative navigation problem for a mobile Navigation Controller. For communities of mixed navigation equipment types, however, such as Doppler and inertial combinations, the relative errors are not representable by the error models of only one or the other type of navigation. The approach to the mixed-community is to prescribe an augmented Rel Nav filter which includes additional error models to properly represent the composite relative navigation error.

***

E. Geographic Updates

With this implementation the mixed-community can achieve accurate relative position, velocity and heading, similar to the homogeneous community. However, the presence of uncalibrated Nav Controller velocity errors (grid drift) serves to degrade the geographic accuracy of specifically targeted points, regardless of whether the relative navigation problem has been optimized. To address this problem, updates are permitted and processed in both the relative grid and the geodetic frames. Geographic updates serve to improve position and minimize the grid drift phenomenon.

Geographic updates directly improve the grid element processing them. Through this grid element, they are then disseminated to the rest of the community. Other community members may utilize these geodetic sources by two possible means. One procedure is to process a TOA which corrects geo position; the other method is via an offset geodetic update. The offset geodetic update allows mixing grid position data and geo position data to derive a geo position estimate.

One of the prospective users of this reported geo information could be the Navigaion Controller, itself. When the Nav Controller processes geodetic data it serves to directly calibrate the rate of grid drift. Successive iterations to other geo sources serves to further refine drift calibration accuracy. Grid and geodetic updates to position are separated to the extent that corrections made to the position coordinates in one frame do not couple directly to the other frame. What this separation provides is the ability to sustain a tactical exercise with smoothly varying grid coordinates unaffected by possibly large perturbations which could be induced when processing geographic updates.

F. Interplatform Error Propagation

An essential consideration in providing accurate relative navigation (in the absence of ground-fixed references) is the inheritance, or correlation, of errors between the community members and the Navigation Controller. The ranging capability of the system provides good instantaneous relative position information; but to achieve accurate long term relative navigation during periods of poor community geometry and ranging signal dropout, the system's Kalman filter should have an estimate of relative velocity errors and relative sensor-cluster attitude errors. In this manner, grid position will be time propagated with the same Schuler error characteristics in each system.

Consider the two aircraft cases where both vehicles are equipped with inertial dead-reckoners. In predicting the range measurement, the filter utilizes the estimated own platform grid position and the reported grid position from the observed aircraft. The range residual will be a function of the TDMA system error, the own platform position estimate, and the observed aircraft position estimate. The net range difference (predicted-observed), regardless of which system position estimate is "right" or "wrong", is used to drive the Kalman filter.

Since both systems present the same error model and same error time signatures, the filter cannot distinguish between errors due to the Navigation Controller's inertial system and the on-board inertial system. This means that the filter recovers the error differences, and hence the relative position, velocity, and INS errors with respect to the Navigation Controller. (By established protocol the Nav Controller will not use grid observations to other members.) Thus the User inherits the errors of the Controller in order to be able to maintain accurate relative navigation. Unlike the Schuler oscillation errors, dynamic effects such as the trajectory-dependent pointing error causes degraded performance in the inheritance process. With good quality inertial systems, however, this is not a major problem.

When additional information is available to the User, such as geographically derived observations external to the grid, separation of User errors from Controller errors becomes possible. If not properly implemented, this will cause decorrelation of errors and hence degraded relative navigation. To obtain the benefits of the geographic update and still maintain provisions for accurate relative navigation, the filter state vector must be augmented with Controller error states. Error sources which are strongly time correlated (e.g. Schuler oscillations), rather than trajectory correlated (pointing error), can be effectively accounted for in this way.

When the two aircraft are equipped with different types of dead-reckoners with different error signatures, the Controller and User errors can be separated without externally derived geographic data. Again, the dominant error sources of the Controller's system must be modelled in the User's Kalman filter for accurate long-term relative navigation to be possible.

Having introduced the concept of the User being able to separate, and thus recover, the controller navigation errors, it would seem advantageous to develop a mechanism whereby the Controller could be informed of this information (without disturbing the error correlations essential to accurate relative navigation). The dual grid navigation approach provides such a mechanism.

G. *Usage of Relative and Geographic Grids

Of utmost importance is the seemingly complex decision as to which grid to use for any given problem. Without geographic observations, the relative grid origin may not be accurately known in terms of geodetic position. Consequently, the use of grid position to locate a point known accurately in geographic coordinates may not result in best performance. A point defined by a source outside the community, such as one given by map coordinates, may best be located using the system's geodetic position estimates. A position identified by another community member using his on-board sensors and grid navigation would obviously best be located using grid position.

The concept of having to select from two different sources of position is not new to the weapon delivery software designer. Most tactical aircraft are equipped with both a radar altimeter and a baro-altimeter, one providing relative (ground referenced) altitude, the other providing a measurement of absolute (sea level referenced) altitude. The use of two altimeters with the associated question of which to use at any given time is well understood and the application-related solution taken for granted. Use of the dual grid data is no more complex than this. In a direct air-to-ground attack of a target acquired by the aircraft's own radar, the software establishes its own local relative grid for navigation to determine when to release the weapon. Since it is a relative problem, use of geographic position directly is meaningless. In addition, any updates to the geodetic system received during that time would be ignored insofar as the position computation is concerned. Over the short period of time (approx. 5 to 10 seconds) from target designation to weapon release, the integral of dead-reckoner velocity, with its inherent short term stability, represents the best estimate of position change.

This single aircraft example is analogous to the community problem where the target is acquired using the radar on one member's aircraft, and this target must subsequently be attacked by another member aircraft. Using the relative grid coordinates, the first aircraft can accurately transmit the location of the target to the second aircraft, just as though the radar had been mounted on that second aircraft. The second aircraft can then attack the target continuing to navigate using the same common relative grid. Any geodetic navigation position observations processed during this time will not impact the relative navigation solution. This is facilitated by the dual grid capability afforded by the terminal where navigation is essentially independent between the two coordinate systems or grids. Updates to the geographic system will correct the geographic position, but will not cause short term decorrelation of position data in the relative grid frame.

*Special section contributed by W. Steele, Singer-Kearfott Corporation

A second example, again beginning with the conventional single aircraft case, can be utilized to demonstrate when position data from the geographic frame would yield better results. A target is designated by a forward observer using geographic map coordinates. An attack on such a geographically designated target should be made based on the aircraft's best available geographic position. All geographic position updates available prior to the attack would be utilized to improve his estimate of geographic position and compute a more accurate weapon release point.

Similarly, if this forward observer information is received by a community member equipped with the dual grid navigation capability of the JTIDS terminal, that member aircraft would also use his best estimate of geographic position. By utilizing data from the geographic grid, he automatically receives the community's best estimate of geographic position, which has been derived using essentially all geographic update information available to each of the community members.

***

H. Alternative Organizational Concepts

The JTIDS software incorporates the concept of a Nav Controller and a grid in which other members are responsible for computing their own navigation states. As such, the current JTIDS can be classified as a "Republican" type of organization.* It is possible to formulate many alternative Republican mechanizations by modifying the definition of hierarchy through protocol and source selection routine changes. For example, source selection could be made more dependent on geometry. The rate at which other members TOA's are used could be made proportional to reported quality numbers, rather than choosing the best over a short cycle of time. Thus, some less accurate members would be given an occasional vote, distributing the navigation responsibility. Pointing as well as position quality could be used for defining hierarchy since pointing is the most difficult yet important error to minimize.

The concepts currently incorporated in the JTIDS software have been studied and been found to be suitable. Although there is continued study of alternative Republican mechanizations, changes would be made only when there is unambiguous justification that another mechanization can provide greater accuracy with little or no addition in computational burden.

Other organizational alternatives have been evaluated with respect to the Republican concept and found (for reasons stated below) generally inferior. However, some features of these alternatives have been incorporated into JTIDS software.

One organization is called "Oracular". Here all members of the community use an oracle (such as GPS, LORAN, OMEGA) to establish the positions which they report. The advantage of this concept is that there is a stable grid origin and direction. However, there are three major drawbacks: (1) No oracle now exists worldwide with enough precision to make it competitive with relative ranging concepts; (2) the oracle would be external to the community and thus vulnerable to attack and to control by forces other than the community; (3) even when such an oracle is universally available (e.g., GPS) it could be cost effective to still have lower capability members or expendable community members who navigate with relative ranging concepts (inherent in JTIDS). Note that JTIDS software has the potential to incorporate oracle information through its geo update features. Thus as GPS becomes operational, its navigation capability can be integrated into the community.

Another organizational concept is called "Authoritarian". Here one member of the community, the master, is responsible for computing the navigation states of all members of the community either through its own measurements and reported dead reckoning information and/or through reported measurements. With all the community navigation information in one location there is the potential for nearly optimal navigation algorithms: the commander would have the complete navigation situation under his control. However, there are several drawbacks: (1) the computational burden on the master would be enormous in a large community; (2) allocation of communication space for navigation would have to be increased dramatically in order to receive reported navigation measurement information and for the master to transmit the navigation states of those community members who require it; (3) loss of master would mean total collapse of the grid. A corollary to this is that only specialized platforms would have the capability of being master. The Authoritarian concept, however, may be used as a grid substructure when guiding RPV's and other low capability platforms to target.

A third organizational alternative, one which is quite appealing, is the "Democratic" organization. This is similar to Republican concepts except there is no controller, or the function of the controller is assigned by the community at large according to real time information. Such members select and weigh others' measurements according to some predefined protocol. Here the loss of a given member would have little impact on the navigation capability of the remaining members. In addition, this organization has the potential of (via averaging) reducing grid drift as the community size increases. Unfortunately simulation studies (4) of various mechanizations of the Democratic grid have failed to demonstrate this potential. Typically a mechanization would have improved grid drift characteristics, but poor relative navigation qualities. Those mechanizations which had relative navigation qualities approaching (but not reaching) that of the Republican concepts suffered from the same grid drifts found in the Republican organization concepts. Note however that JTIDS source selection routines are such that the community effectively reverts to a Democratic organization when a master is lost and not yet replaced. Accuracy does degrade, but some semblance of relative navigation is maintained.

* Republican organization - master selected from peer group.

VII. HYBRID NAVIGATION DATA PROCESSING

W. J. Steele, Singer-Kearfott Corporation

The navigation software package developed for JTIDS serves to integrate the community and the on-board navigation resources to provide the Grid Acquisition and Grid Update capabilities. This section of the paper is devoted to a functional description of the basic software modules with special emphasis on the Kalman filter.

***

A. Functional Summary

In conjunction with its interfaces with various terminal processor I/O ports, the following specific tasks are performed by the navigation software:

Continuous time base synchronization control in both active and passive synchronization modes.

AN/ASN-91 navigation port interface processing.

Relative navigation grid acquisition and update.

Inertial integrated Kalman filter (hybrid) navigation in both active (RTT) and passive synchronization modes.

Mode control

The JTIDS computer provides control for each relative navigation mode of operation. The hybrid navigation processing provides navigation data in both relative grid and geographic coordinates and accepts data from the dead reckoning system navigation port. The relative navigation function is dependent upon various message processing and time-of-arrival (TOA) measurements and message data containing the relative grid and geographic position of other community members.

B. Program Interface with the AN/ASN-91 Computer

The terminal processor interfaces with the AN/ASN-91 Weapon Delivery Computer for the following purposes:

1. To provide the relative navigation program with periodic dead reckoning inertial navigation variables and time mark.

2. To provide the AN/ASN-91 weapon delivery computer with relative navigation grid position and estimates of inertial navigation system errors.

The timing of the data exchange is controlled by the AN/ASN-91 computer, and this timing is asynchronous to JTIDS timing. To achieve synchronous operation of the two computers, the AN/ASN-91 computer provides a time mark interrupt indicating the time at which a batch of data provided is valid. One of the words provided by the AN/ASN-91 contains the AN/ASN-91 time word corresponding to the time of the interrupt. Thus a set of AN/ASN-91 interrupts and time words is sufficient to define the AN/ASN-91 computer's time base. In addition, one of the words in the data batch transferred to the AN/ASN-91 time tags the JTIDS data to the AN/ASN-91 with respect to the AN/ASN-91 computer's time base.

C. RTT Synchronization Filter

The RTT Synchronization Processing function is used to mechanize the fine oscillator synchronization in the RTT Synchronization mode, following confirmation of coarse synchronization. Synchronization of the oscillator is achieved by estimation of the oscillator phase error (ω) and frequency error (θ) relative to the reference time base. The estimator is mechanized as a two state Kalman filter, modeling the previous states, and using direct observation of oscillator phase error supplied by the RTT Reply Processing routines. The phase and frequency errors thus determined by the filter are used by the Time Base Control Processing to generate correction commands to the oscillator. Corrections to the phase and frequency computed by the Time Base Control and accompanied by corresponding decrements of the Kalman filter state estimates.

D. Relative Navigation Source Selection and Message Processing

The Relative Navigation Source Selection and Message Processing function screens all validly received P-messages for possible sources for use as either geographic updates or relative grid TOA updates. The screening criteria is based upon own-unit position, time, and azimuth quality data and own-unit user designations. User designations include Navigation Controller, Primary User, Secondary User, and Position Reference. Sources are screened for time and position qualities superior to corresponding own-unit quality levels. This community source selection procedure time tags selected source TOA data and message data against a dead reckoner counter and saves corresponding own unit navigation data for use by the Kalman Filter Processing function.

E. Relative Navigation Port Interface

The Navigation Port Interface Processing function provides the computations necessary to interface input and output data between the Kalman filter processing and the dead reckoning navigation port. This interface is compatible with an inertial dead reckoning subsystem. The proper buffering control and timing is supplied to allow operation with a dead reckoning subsystem supplying data referenced to a time base which is independent and asynchronous to the slot interrupt times. This processing is capable of operating with dead reckoning input data periods of 250 ms.

F. Kalman Filter

The Kalman Filter Processing routines uses selected P-message observation data with dead reckoning port input data in a filter algorithm which estimates relative and geodetic navigation position, velocity, and inertial platform attitude errors. This filter algorithm shall be capable of operating in RTT synchronization or passive synchronization modes. Relative navigation grid definition and acquisition computations shall also be performed.

The Kalman filter is the heart of the relative navigation data processing function. A fifteen state filter is employed in the current system wherein the ranges to other members, as derived from the time synchronized ranging capability of the Time Division Multiple Access (TDMA) communication system, are used as the observables to the filter while velocity information from the dead reckoner allows the state vector to be time propaged between observations. Currently configured for use with an inertial dead reckoner, the filter state vector includes elements for North and East velocity, altitude, latitude and longitude, three IMU attitude errors, two components of grid position, grid azimuth angle with respect to North, and the TDMA system clock bias and frequency errors. The filter could be easily modified to accomodate other types of dead reckoning or navigation equipment. This has been done for a Doppler/Attitude and Heading Reference System (D/AHRS) in a separate demonstration for the Navy.

During intervals between observations, the velocity data from the dead reckoner are utilized in the differential equations for the states to propagate them forward in time. When it becomes time to process an observation, the software Source Selection Routine searches a buffer of valid range measurements to find the ones which provide the most beneficial geometry and highest quality position data. The extrapolated grid position is time correlated to the TDMA slot of the observed member. Using the reported grid position from that member and the estimated grid position for the user aircraft, a predicted range is determined. The difference between the predicted range and measured range provides information concerning errors in the state vector elements. The Kalman gain correctly partitions this range a residual among the Kalman states. Nominally, three range observations are processed every 16 seconds.

Having entered the grid community, each member also transmits his estimated position and quality for others to observe. The qualities represent the system's statistical estimates of its own navigation errors and are assigned based predominantly on the filter's covariance matrix. They are used by other members in their observations selection logic in an effort to utilize the most accurate data. A separate quality is transmitted for geographic navigation accuracy and relative grid navigation accuracy.

References (2) and (5) provide an analytic description of the basic algorithms. Reference (6) provides a comparable discussion of a JTIDS/DAHRS system as developed under the ITNS (Integrated Tactical Navigation System) program.

G. Other Relative Navigation Software Modules

The Relative Navigation Input Data Display Processing function provides the man-machine interface for relative navigation input data and status variables. This module formats data for presentation on the control/display unit and accepts operator inputs. This module interfaces with the Kalman Filter module.

The Relative Navigation Output Display Processing function provides a set of navigation parameters for display on the control/display unit.

The Message Handler Filters for Relative Navigation Source Selection and Message Processing provide preliminary screening and set up of received position reports. These modules are part of the received message handlers and interface with P-message storage buffers.

The Relative Navigation Net Entry Control Channel function is used to support the transmission of the Net Entry Control Terminal's Net Entry Signals, Position Reports, and RTT Replies.

The Message Handler Filters for RTT Synchronization Filter Processing function provide preliminary screening and set up of received position reports in order to identify potential candidates for RTT interrogations. These modules interface with the P-message storage buffers.

The RTT Interrogation Generation function formulates RTT interrogation messages when sources having greater time quality than own-time quality are found by the P-Message Handler Filters for the RTT Synchronization Filter. The RTT interrogation messages are therefore formulated for transmission at variable rate and only when required. The System Reference Number of the selected source is formatted within the RTT interrogation messages.

The RTT Reply Processing modules are used to collect and process data from RTT replies stored in the received message buffers. This data shall be provided to the RTT Synchronization Filter.

The Time Base Control function is used to accept either RTT Synchronization Filter or Kalman Filter oscillator phase and frequency corrections and forward these requests to the time base in the Signal Processor.

VIII. SUMMARY

Lawrence Newman, Naval Air Development Center

JTIDS provides the power of instantaneous communications, digital data processing and relative navigation to maximize the integration and coordination of available resources - land, sea and air. Our discussion centered on the relative navigation function wherein we have endeavored to explore the objectives and generic features of relative navigation, we have looked at the ADM product, and we have examined some of the many potential applications. The task which lies before us now, as the program advances from ADM to EDM to a full fleet capability, must be to bridge the gap between concept and capability and bring to fruition the full measure of the system's potential.

IX. REFERENCES

1. Brentnall, Col. B., USAF, "The Joint Tactical Information Distribution System," NATO AGARD Symposium, FRG, 1977.

2. R. C. Stowe, Singer-Kearfott, "Tactical Hybrid Navigation Scheme for Mixed Community Coordination," IEEE EASCON 1973, September 1973

3. Bauer, Mason, and Wilson; M.I.T., "Radio Refraction in a Cool Exponential Atmosphere," Lincoln Tech Report 186; 27 August 1958.

4. Rome, Dr. H.J., Dynamics Research Corporation, "Evalution of the Accuracy of Community Relative Navigation Organization Concepts," ION National Aerospace Symposium, 27-28 April 1976.

5. Danik, Dr. B., Singer-Kearfott Corporation, "Hybrid-Inertial Navigation with Range Updates in a Relative Grid," AIAA Guidance and Control Conference, 20-22 Aug 1973.

6. Bivin, James D., Naval Air Development Center, and Solowey, Greg, Singer-Kearfott Corporation, "Testing of an ITNS/DAHRS System," ION National Aerospace Meeting, 12-13 May 1975.

7. Rubin, Jack, ITT, "DTDMA in JTIDS," Military Electronics/Countermeasures, Oct 1977.

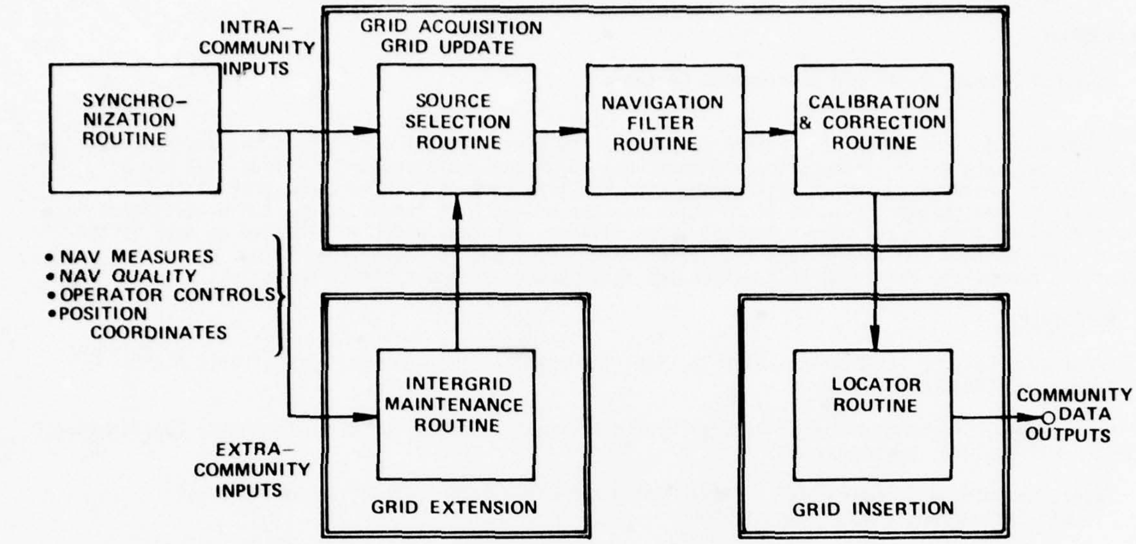


FIGURE 1 - JTIDS Community Navigation Net Functions

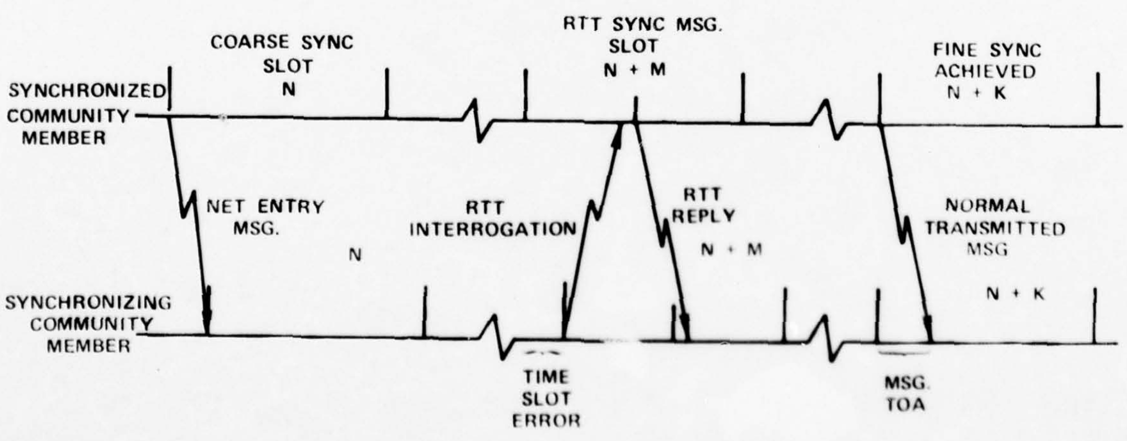


FIGURE 2 - TDMA Slot Time Synchronization

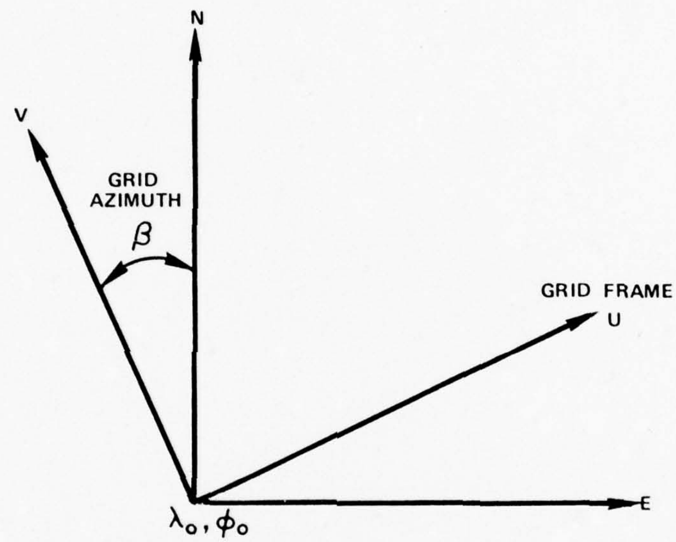


FIGURE 3 - Grid Coordinates

INTEGRATED TACTICAL NAVIGATION SYSTEMS (ITNS)

EDWARD DINTER
SPVY NAVIGATIONAL SYSTEMS ENGINEER
JAMES D. BIVIN
ELECTRONICS ENGINEER
NAVAL AIR DEVELOPMENT CENTER
WARMINSTER, PA 18974

SUMMARY

The relative navigation process used in JTIDS was developed and tested under the Navy sponsored ITNS program. Over a four year period, from 1972-1975, a number of flight tests and air/sea demonstrations were conducted to explore the capabilities of this new navigation concept. Although many variations have been addressed, the basic principles remain unchanged, and the operating experience gained with ITNS is applicable to JTIDS. This paper describes the five demonstrations that were performed over the period and their associated results.

V. A.2 INTEGRATED TACTICAL NAVIGATION SYSTEMS (ITNS)

A. BACKGROUND

ITNS combined range measurements between vehicles with self-contained dead reckoning in each vehicle to achieve accurate tactical navigation in spite of geographic navigation error. A navigation grid was established arbitrarily and grid position estimates from the dead reckoning were reported via data link. Range between vehicles was measured and compared with the expected range as calculated from the reported position, thus determining the error in position estimate. Corrections were applied to the dead reckoning and after several minutes of operation the dead reckoning would be aligned with the grid. Vehicles located themselves in this manner independently and continuously, and the effect was to minimize the relative error between vehicles.

The ranging and communication functions were served by an experimental TDMA data link known as TSRS (Time Synchronized Ranging System). This was a fixed frequency, time-multiplexed system in which each terminal had an assigned time slot for transmission and received in all other time slots. User terminals synchronized independently with a master terminal to a high degree of precision, and TOA (Time of Arrival) of transmissions was a measure of range to the transmitter. The TSRS had an active (transmit/receive) and a passive (receive only) operating mode. In the active mode, synchronization was automatic and TOA measurements were used as range measurements to estimate grid position. In the passive mode, TOA measurements included a clock bias which was estimated by the computer along with the grid position.

The TSRS, although less sophisticated, was similar to JTIDS in many respects. It operated in the L-band at 1140 MHz using differential phase shift keying with 13 bit Barker coded synchronizing signals. All transceivers were identical and included automatic coarse and fine synchronization with an arbitrarily selected marker. (The equivalent of the JTIDS Net Time Reference). There were 1024 user slots every 2.6 seconds in which a complete epoch of synchronization, range measurement between all members, and transmission of each member's grid position occurred.

The ITNS program was a concept development rather than a hardware development and available equipment was used wherever possible to assemble the prototype systems. The development effort concentrated on the multiple vehicle system concepts and on the real-time computer program to implement these concepts. The TSRS had been developed by the contractor on IR&D funds, and three terminals were purchased and three borrowed for use in flight testing. Inertial navigation systems from government inventory were utilized for the dead reckoning function.

ITNS was first proposed by the Singer-Kearfott Company to Naval Air Systems Command (NAVAIRSYSCOM) in early 1971, and Naval Air Development Center (NAVAIRDEVCEN) was tasked with conducting the development. A contract was let with Singer and in May 1972, two aircraft systems with spares were delivered. The following paragraphs describe the five tests and demonstrations performed with this equipment to confirm the feasibility and the utility of the ITNS concept.

B. DEMONSTRATION DESCRIPTIONS

1. AUTEC (ATLANTIC UNDERWATER TEST & EVALUATION CENTER)

a. Purpose

The purpose of this test was to ascertain the feasibility of the ITNS Concept. It should be noted that the ITNS project was a concept development, not a hardware development and off-the-shelf equipment was used to assemble the prototype system.

b. Description

From September 1972 to May 1973, flight tests were conducted over the AUTEC instrumented test range at Andros Islands, Bahamas.

The AUTEC test range, although primarily an in-water tracking range, has an in-air tracking capability consisting of two precision radars and five cine-theodolites. During testing, one radar would track each aircraft and provide a position history on magnetic tape. Meanwhile,

aboard each aircraft, a history of the ITNS position estimate was recorded on magnetic tape. After the test, these tapes were time-correlated and compared to determine the position accuracy of the system. The aircraft trajectories were designed to stay fairly close to the radar sites in order to maintain tracking accuracy. Under these conditions, it was expected that the radar accuracy would be sufficient; and theodolite data, while collected in limited quantities, was therefore not processed.

The aircraft were based at Key West, Florida, approximately 200 miles from the range. During a test day they would fly to the range, conduct several exercises in succession, and return to base. Test coordination was provided by the AUTEC test director who monitored aircraft position, instrumentation status, and progress of the exercise. Exercises varied from a few minutes to several hours in length, but most from 30 minutes to 1 hour. A total of 198 exercises were conducted in 43 test days during the eight-month period.

The three ground stations were established at fixed locations on the range for the duration of the test series. These ground stations were used to simulate additional members of the community, although they had no dead reckoning or computational capability. Because they were stationary, their position in the grid could be determined from the aircraft by a trilateration procedure, and this position was used as if it were a report from a full-fledged member. Different community sizes were evaluated by using all, some, or none of the ground stations in addition to the two aircraft. A tape recorder at one of the ground stations recorded the position reports of the aircraft as backup instrumentation.

Large aircraft were required for use as test beds to carry ITNS equipment, instrumentation, and technical personnel while remaining airborne for extended periods. Three different aircraft were used during the course of the test: a P3 long-range patrol aircraft; a C117 cargo aircraft during the 1972 test; and the same P3 along with an EC121 long-range patrol aircraft during the 1973 testing. During most of the exercises, the aircraft flew racetrack patterns at altitudes of 7,000 to 15,000 feet and speeds of 150 to 250 knots. Some trajectories were designed to study the effects of different geometry but tactical flight patterns were avoided since they caused loss of radar tracking.

c. Results

Feasibility - The ITNS concept, combining range measurements between vehicles with self-contained dead reckoning, was found to be feasible. Communities of two to five members were attempted, and in each case stable navigation was achieved. Furthermore, the navigation accuracy, was sufficient for most tactical applications. The capability for some members to participate without radiating was particularly appealing. In-air alignment of inertial systems was accomplished with ITNS. Fixed ground stations were not required, but performance degraded in communities smaller than four members.

Applicability - ITNS was found to be an effective utilization of the TOA measurements from a synchronized TDMA data link. Given a community of vehicles equipped with dead reckoning and TDMA, the techniques of implementation were simple and straightforward. ITNS and TDMA complemented one another, and a large payoff could be realized at small additional cost.

System Concepts - The flight testing disclosed the following aspects.

Grid Acquisition - Two methods of grid acquisition were tested: trilateration and inertial. After a few minutes of grid operation, performance was equivalent. Either method was suitable although inertial acquisition was simpler.

Reset Capability - As a result of necessary approximations, the Kalman filter would occasionally become nonlinear and a recovery method for such situations was necessary. Maintaining free inertial dead reckoning, as well as the filter-corrected hybrid navigation, serves this purpose and was well worth the additional complexity.

Reasonableness Tests - Checks for consistency of range measurements and dead reckoning detected occurrence of improper operation. Such tests gave warning of unreliable operation and indicated when a reset was required.

Stationary or Slow Moving Members - The inclusion of a single stationary or slow moving member in a community can cause accuracy to degrade. The member can locate himself properly but does not make a good reference for faster moving members. Stationary or slow moving members should either be used in pairs or else assigned a secondary standing in the community.

2. AT-SEA FEASIBILITY

a. Purpose

This demonstration was conducted for the purpose of defining the operational feasibility and potential tactical utility of the ITNS system concept, showing that the results of the AUTEC test could be extrapolated to tactical situations.

b. Description

Airborne Functional Description

The basic airborne unit configuration consisted of the experimental ITNS flight system flight tested at AUTEC in two aircraft. Two such systems were installed in aircraft, one in a P3A and the other in a C121, for the USS GUAM demonstration.

A capability to transmit and receive target position data in community grid coordinates was added. Manual insertion of target range and relative bearing, together with a time mark, enabled the tracking (transmitting) aircraft to calculate target grid position. This data and other manually inserted data was transmitted to all members of the ITNS community.

A capability was also added so that any receiving aircraft could continuously compute and display own range and true bearing either to a target received from another vehicle via the ITNS data link or to grid coordinates inserted manually by the operator

Shipboard Functional Description

The USS Guam was configured as a full ITNS community member. The hardware configuration was similar to the airborne inertial configuration.

The major difference in the shipboard configuration was that the ship's Electromagnetic (EM) log provided speed to the digital computer. It was necessary to use ship's speed to damp the inertial system's oscillatory errors which tended to build with time. Since the inertial set was designed for aircraft flight times of a few hours, it was necessary to provide the damping on a ship where operational periods are measured in days or weeks. A shipboard OMEGA navigation receiver was installed on the USS GUAM. Position fixes from this receiver were used to correct the drift of the ships ITNS inertial set.

For the shipboard command and control functions, the grid position of each ITNS community member and all reported targets were plotted on a surface grid plot. The dynamic tactical situation was thus immediately available. For the demonstration, grid coordinates were also transmitted via sound-powered phones to the combat Information Center where an extended surface plot was maintained. The ITNS on board the ship was also provided with the computer capability to calculate the range and bearing to any community member, any transmitted target, or any grid coordinate inserted manually into the computer.

The shipboard system was capable of operating in all the modes of which the airborne units were capable, including the passive mode, so long as two other active community members were within LOS (line of sight) of the ship. Additional modes were available for experimentation as outlined in Table I.

TABLE I

ITNS OPERATING MODES

Mode I:	Airborne controller, airborne user, marine user
Mode II:	Marine controller, airborne users
Mode III:	Airborne user, 1-3 ground stations (not used in USS GUAM demonstration)
Mode IV:	Airborne controller, airborne alignment of user inertial system for subsequent geographic navigation, marine user
Mode V:	Mode I with geographic position updates derived from external geodetic reference applied to controller
Mode VI:	Mode I with any user operating passively (no RF emission)
Mode VII:	Same as Mode VI, but with initial active synchronization

Modes I and II include the capability for the airborne user and controller to interchange functions without losing the grid.

Demonstration Phases

The demonstration of the ITNS capability in the deep-ocean environment was designed to indicate some of the potential practical advantages of the precision navigation capability and tactical data exchange potential of the ITNS.

The demonstration was designed around the minimal community of platforms which had been configured as ITNS relative grid community members. This community consisted of the USS GUAM and two aircraft: a P3A and a C121. Because the two aircraft were staged each day from the Naval Air Development Center (NAVAIRDEVCEN), Warminster, Pennsylvania, and because the USS GUAM was operating in coastal waters near Charleston, South Carolina, the demonstration was limited to an aircraft overhead period of approximately 3 1/2 to 4 hours each day.

Because of the difference in speeds of the two aircraft, the P3A usually arrived earlier than the C121, and thus was assigned the function of establishing the common grid and assuming the role of controller.

The demonstration scenario consisted of the following basic phases:

Pre-Demonstration Phase - In a phase preceding the demonstration, the common grid would be established by the P3A and acquired by the other two platforms, and the two aircraft would then take up their positions for the first phase of the exercise.

Phase I - A/C would report targets

Phase II - Demonstrate over-the-horizon capability of the ITNS (using relay A/C)

Phase III - Demonstrate passive operation

On those demonstration occasions when visitors were aboard one of the aircraft instead of the ship, a fourth demonstration phase was conducted. During this phase, the capability of an aircraft to navigate precisely in the relative grid was demonstrated in a passive mode. Station keeping capability between the two aircraft was demonstrated in an active mode, and the total range capability of the ITNS radio system was demonstrated.

c. Results

The demonstration aboard the USS GUAM was conducted in an abbreviated at-sea period from 10 January to 26 January 1973. In preparation for this demonstration, only two ITNS aircraft flights were made in the neighborhood of the USS GUAM; both of these were made during the in-port period in December 1972. A single cruise period in December had been used to check out the at-sea alignment of the inertial system, using hand-set EM log velocity reference and OMEGA position corrections. These tests for USS GUAM system check-out at sea were made without the presence of the ITNS aircraft.

The first community operation of the ITNS relative grid system using all moving vehicles in the open-ocean environment was accomplished on the very first flight of the two test aircraft in the neighborhood of the at sea USS GUAM. On this date, 11 January 1973, the ITNS demonstrated the capability to establish a common grid, the capability for all units to acquire that grid, and the capability to continuously navigate precisely in the common grid. Continuous position reporting by the ITNS data link was demonstrated, and the initial use of tactical commands using the command library was demonstrated. Also on 11 January the passive operator mode was demonstrated.

During the succeeding days of the at-sea period, additional items of the at-sea demonstration scenario were successfully added to the schedule of each given day until the full demonstration capability could be exercised in a single 4-hour on-station period.

On 17 January 1973, a full demonstration scenario was conducted. All elements of the demonstration scenario were again accomplished successfully. The scenario rehearsal was performed in preparation for a full scale ITNS capabilities demonstration which was conducted for flag rank officers and civilians of the Navy Department who observed the full operations aboard the USS GUAM on 18 January 1973. Again, during this formal demonstration all demonstration capabilities were successfully accomplished with the demonstration scenario.

An additional formal demonstration was conducted on 23 January 1973 for Navy, Air Force, and civilian observers. These observers witnessed the demonstration from the vantage point of the P3A. During this day and this day alone, the controller function was assumed by the C121. This was done in order that the visitors might observe common grid acquisition by a user from the platform on which they were riding.

Again, all items of the demonstration scenario were successfully accomplished. Passive aircraft operation was successfully demonstrated at that time. In addition, the station keeping capability of the two aircraft was demonstrated by assuming a close formation flight. In this maneuver, the two aircraft were approximately 350 feet apart, with the P3A assuming a position 250 feet to the rear and 250 feet to the starboard of the C121. The system consistently indicated the proximity of the two aircraft during this maneuver. The range test was also conducted to demonstrate the reception range of the ITNS data link and the DME.

3. TACTICAL/SEA CONTROL SHIP

a. Purpose

The purpose of the demonstration was two fold:

Ascertain the benefits of ITNS to the SCS (Sea Control Ships) operations

Develop potential new tactics using the increased capabilities offered by the ITNS.

b. Description

The ITNS was modified for use during this sea test demonstration. This demonstration system operated as part of a community of users consisting of three members, including a shipboard system and two helicopter systems. This modified system had the following capabilities:

It was designed to provide precision navigation to users within a community relative grid. Through the proper use of this navigation grid, targets could be accurately referenced to all community members.

A line-of-sight data link communication capability was provided whereby target data and message data could be transmitted between any or all community members. Each member's grid position and status was automatically transmitted over this link while in active mode.

An automatic data relay was provided whereby a controller designated unit (a capability which can be assumed by any member) would automatically relay the data between users not within line-of-sight of each other. Any user could also operate within the navigation grid and receive all data link communications while in passive mode (no RF radiation).

Prior to in-flight grid navigation, deck initialization and alignment of the airborne inertial systems to the ship's ITNS inertial system was provided through use of the same radio data link described above. ITNS grid operation was exercised by an airborne unit while proceeding with its deck initialization and alignment.

Following initialization and start up of the airborne or ship systems, the focal point of operator mode control and tactical coordination was primarily centered about two control and display panels. These were designated as the CD (Computer Display) local TDP (Tactical Display Panel). In addition remote display of navigation and mission data was supplied in the helicopter configurations through an additional "display only" remote TDP and BDHI (Bearing Distance Heading Indicator) at the pilot's station in each aircraft. Also, two additional remote "display only" TDP's were supplied on the ship, with one located in the CIC (Central Information Center) and one located in the ASCAC (Anti-Submarine Classification and Analysis Center).

Operator navigation mode control of the system aboard ship or in either helicopter was exercised through use of the CP (Control Panel). This panel permitted the operator to command ITNS grid navigation following alignment and to command normal geographic navigation modes. All tactical mission operations were exercised through use of the primary TDP. Passive mode operation was commanded by a user community member from this panel. Both target location options (flyover and offset) were initiated through the TDP with an ability to command storage of up to 10 targets. Range, relative bearing, and time-to-go to any of these 10 stored targets, as well as both other community members, were displayed on the TDP by operator switch command. This display was used in the helicopters for vectoring to targets positionined in the ITNS grid or to other community members. On board the ship, this display capability was used for plotting grid positions of received targets and the two helicopters relative to the ship on the tactical plotting board in the CIC. Target and message data transmissions, as well as acknowledgements of received transmissions from other members were initiated through the TDP. All communication traffic could be addressed to one or all community members.

In the helicopters, the remote TDP display at the pilot's station provided aircraft velocity and true heading as well as the same target data displayed on the local TDP. This, together with the BDHI display, provided the pilot with pertinent ITNS navigation data and targeting data for use in vectoring to various target positions.

This same remote TDP capability in the CIC and the ASCAC on board the ship allowed position location and status of both aircraft and targets to be monitored.

c. Results

Comparison of radar and ITNS position marks indicated a difference greater than that which could be associated with radar error. The difference was due to two sources; EM log error and ITNS grid drift. This grid drift is consistent for all platforms utilizing ITNS, and thus does not affect tactical decisions made based on grid data.

This comparison indicated that the velocity-heading dead reckoning should be part of the ITNS for sea control ship operations.

The objective of the exercise was to detect and locate a participating submarine. During the first four days of the exercise, the ITNS was used sparingly, mostly for data transmission of buoy drop commands. Positions of aircraft were marked on the DRT (Dead Reckoning Trace) by reference to radar, rather than ITNS. Commencing on 12 September 1973, the DRT positions were marked by the ITNS, and checked by radar for operator confidence.

The success of the ITNS was best demonstrated by the comparison of using different new methods for laying sonobuoy barriers. To have a high degree of confidence the barrier spacing must be less than a certain distance apart. The barrier was consistently penetrated until it was layed utilizing the ITNS.

One of the major contributions of ITNS was the reduction of voice communication. In preparation for this tactical demonstration, the helicopter crews developed a standard set of messages to be exchanged using the ITNS data link. This set of messages was changed and increased as the demonstration progressed and as more uses of the data link were developed and implemented by shipboard and helicopter personnel.

The first day of the ITNS demonstration both helicopters were sent out to perform over-the-horizon relay and target hand-off functions. One helicopter went over-the-horizon, marked a target, and sent back through the relay helicopter the target position. The over-the-horizon helicopters' positions were also continuously and automatically relayed to the USS GUAM through the high flying helicopter. Target hand-off between the helicopters was then performed just by turning a switch and pushing a button. No problems were encountered in the relay or hand-off function demonstration. It should again be noted that the target information and the position of both helicopters were available in the USS GUAM's CIC via the remote TDP.

On most days of the demonstration only one ITNS equipped helicopter operated with the ship during a particular exercise. Although operations involving more than two platforms provide the most adequate mode of ITNS operation, the one-on-one operation proved very satisfactory. The barriers were layed, in fact, using this one-on-one technique.

This performance level was maintained while experimenting with various types of barriers during the remaining days of the exercise. Methods were also developed by helicopter personnel to reduce the time to lay various patterns by using ITNS. The helicopter personnel were the prime developers for these new tactics and were more than satisfied with the ITNS.

4. ITNS/CONDOR DEMONSTRATION

a. Purpose

The purpose of the ITNS/Condor demonstration was to show the capability of the radar and electro-optical sensors of the Condor missile, when augmented with the ITNS position reporting and vectoring capability, to accomplish remote surveillance and control. The Condor missile itself was not used in the demonstration, only a pod containing the sensors.

b. Description

The Condor missile contains, in addition to motor and warhead, advanced radar and electro-optical sensors and a two-way data link. After launch, the operator monitors the radar or television imagery returned by the data link and controls the course of the missile to bring a desired target to the center of the display. The target is then designated to the missile which homes on contrast differences in the display, providing great accuracy against fixed or moving targets.

The radar has high range resolution and wide band frequency agility for sea clutter suppression, accurate target designation, and stable target tracking. It has a high area revolution mode for recognition of complex targets and can operate in either monopulse or track while map modes. The electro-optical seeker is a TV camera slaved to the radar tracking line. If a target is designated, it provides automatic lock-on and aimpoint update using contrast differences within the TV tracking gate. The data link provides man-in-the-loop control up to the point of target designation.

The ITNS equipment used in this demonstration was essentially the same as in the previous demonstrations; providing a common community tactical grid for precision relative navigation, position reporting, station keeping, vectoring, and shared geographic navigation. A CRT was included to show a horizontal situation display of the position and identification of moving vehicles and designated points of interest.

The demonstration community consisted of the USS Guadalcanal which served as command and control center, an ITNS equipped CH53 helicopter which served as target designator, and a B26 equipped with the ITNS and the Condor sensor pod. The ITNS and Condor display consoles were installed side by side aboard the ship. In the B26 the missile guidance commands and ITNS vectoring commands were displayed to the pilot who steered accordingly.

The demonstration scenario started with a surveillance mission using the telemetered radar and TV imagery along with routine ITNS position reporting to map out areas of tactical interest, controlled remotely by members of the audience. Meanwhile, the helicopters selected a target of opportunity (a passing merchant ship) and marked on top several times to develop a track. The B26 was then accurately vectored to intercept the track within the narrow field of view of the electro-optical sensor alone. The Condor system then homed on the selected target to demonstrate the automatic tracking capability. Finally, the ship switched to radio silence for rendezvous with the aircraft. No Condor imagery was returned but the audience could compare the horizontal situation display with visual observation of the two aircraft to see that navigation accuracy was maintained.

The demonstration was scheduled for April 1974 while the USS Guadalcanal was steaming off the Virginia coast. However, on the day of the demonstration, a freak spring snow storm halted air traffic and prevented visitors from traveling. The demonstration was rescheduled the following week with the ship at pier-side in Norfolk.

c. Results

The demonstration was a very impressive display of surveillance and navigation capability, combining precise position reporting with high quality imagery and the ability to "steer" the sensor. The combination allows rapid assimilation and accurate evaluation of data, enabling a quick response to threats.

5. D/AHRS (DOPPLER/ATTITUDE HEADING REFERENCE SYSTEM)

a. Purpose

The original ITNS equipment used inertial dead reckoning. The objectives of the DAHRS/ITNS project were to determine the feasibility of using alternate dead reckoning equipment, develop a mechanization using alternate equipment, evaluate the performance of this mechanization, and investigate the compatibility with the inertial mechanization. The project was conducted in three phases; a design study phase, an equipment modification phase, and a flight test phase. The first two phases were performed by Singer-Kearfott under contract, and the third phase was performed jointly by the Army and Navy.

b. Description

The airborne configuration selected for implementation was a doppler radar with an inertial system operating in backup mode to simulate an AHRS. The shipboard configuration selected was synchro velocity input (similar to EM log) and the inertial system in backup mode. Both of these configurations were required to be interchangeable with the inertial configuration by reconnecting cables and reloading the computer program. Furthermore, the transmitted message format was required to be compatible with the inertial message format, so that interoperability of the two configurations could be investigated.

The ITNS equipment was installed in three aircraft, an Army UH-1 and a Navy SH3 and C131. The shipboard station was installed in a laboratory at NAVAIRDEVCEN, and ground stations were installed at Ft. Dix and NAS Lakehurst. There was sufficient equipment to operate five of these stations at one time.

Two redundant instrumentation schemes were used. Each aircraft was equipped with a vertical camera to photograph the ground and a horizontal camera to photograph the instrument panel. Upon overflight of a landmark, simultaneous photographs were taken, allowing offset from the landmark to be calculated. This was supplemented with manual data recorded by the operator. In addition, each aircraft was provided with a tape recorder which recorded all significant variables including barometric altitude and range to two ground stations. This provided sufficient information for trilateration position solutions.

During the period from March through July 1975, a total of 283 hours were flown in various configurations, environments, and operating modes. Performance of individual members with respect to ground stations, and of two and three member communities without ground stations were evaluated. The Army aircraft investigated the effect of low level (nap of the earth) flight and the Navy aircraft investigated over water operation.

c. Results

Due to equipment age and heavy usage, the TDMA and AHRS performance was erratic, but these components were not the subject of this evaluation. The relative navigtion concept, which was under evaluation, was found to be generally tolerant to subsystem behavior.

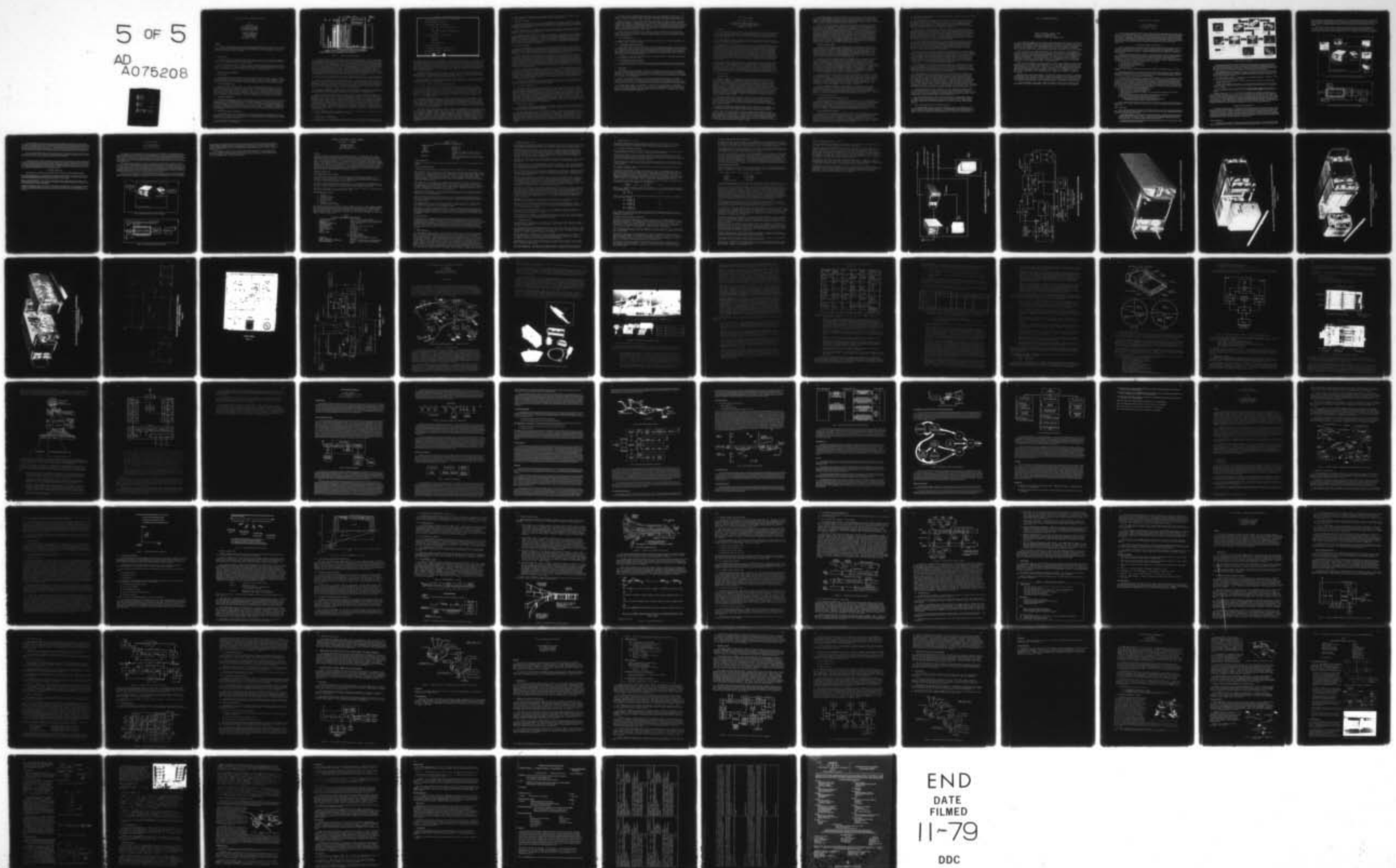
ITNS performance is dependent on geometry. The flight test geometry was neither the best nor the worst possible, and represents a reasonable tactical situation. However, a more widely distributed community would produce somewhat better results. Relative position error was on the order of 3-5 times larger than that obtained using an inertial system as the dead reckoning device. Current studies being performed by the Singer-Kearfott Company (Section III.B) are aimed at reducing these errors in JTIDS configurations.

C. SUMMARY

The orderly progression of conducting demonstrations of the ITNS resulted in wide spread acceptance not only of the feasibility of the concept but in a realization that it could enhance fulfillment of certain missions. The ITNS data base aided the JTIDS development process, especially in the relative navigation area.

AD-A075 208 ADVISORY GROUP FOR AEROSPACE RESEARCH AND DEVELOPMENT--ETC F/G 17/7
PRINCIPLES AND OPERATIONAL ASPECTS OF PRECISION POSITION DETERM--ETC(U)
JUL 79 C T LEONDE
UNCLASSIFIED AGARD-AG-245 NL

5 OF 5
AD A075208



END
DATE
FILMED
11-79
DDC

JTIDS - THE ISSUE OF FREQUENCY SELECTION

BY

JACK RUBIN
ITT AVIONICS DIVISION
CNI SYSTEMS DEPARTMENT
500 WASHINGTON AVENUE
NUTLEY, NEW JERSEY 07110

SVEN H. DODINGTON
AVIONICS CONSULTANT
ITT HEADQUARTERS
320 PARK AVENUE
NEW YORK, N.Y. 10022

SUMMARY

The Joint Tactical Information Distribution System (JTIDS) will provide to the military services improved performance/secure digital communication, precision relative navigation and conventional Tacan and IFF functions. The portion of the electromagnetic spectrum proposed for JTIDS is 960 to 1215 MHz.

1. INTRODUCTION

The Joint Tactical Information Distribution System (JTIDS) is an Integrated Communication Navigation and Identification (ICNI) system that features spread spectrum, frequency hopping, high performance, digital communications.

It is proposed to operate the Joint Tactical Information Distribution System (JTIDS) in the 960 to 1215 MHz Lx-Band, already occupied by the ICAO DME and ATCRBS and by the military Tacan systems. This paper examines the technical and historical reasons for this and shows why compatibility with these existing services is entirely feasible.

2. THE PHYSICS OF THE PROBLEM

2.1 PROPAGATION

For wide-band radio systems operating near the surface of the Earth, there is a broad optimum spectrum from about 200 MHz to about 5000 MHz. Below this band, ionospheric reflection occasionally produces unwanted interference (below 30 MHz it actually becomes a major mode of communication), while above this band, atmospheric precipitation produces losses.

2.2 RECEIVER ANTENNA AREA

Within the 200 to 5000 MHz band, all frequencies propagate equally well, but with significant increases in propagation losses as the frequency increases with the received power being proportional to receiver antenna area. For simple omni-directional antennas, this area decreases as the square of the wavelength; e.g., a dipole at 2000 MHz intercepts one quarter of the power of one at 1000 MHz. Since receiver noise figures, in general, tend to get worse as the frequency is raised, a compensation for reduced receiver antenna area is increased transmitter power. This favors operation at the low end of the band.

2.3 PERCENTAGE BANDWIDTH

Wideband systems become a great deal more practical if they are less than an octave wide as this avoids the second harmonic of a low-frequency channel from being mistaken for the fundamental of a high-frequency channel. A second consideration for avoiding an octave or greater bandwidth is that a quarter-wave at one end of the band is a half-wave at the other end, thereby complicating the design of tuned circuits. Thus, a "comfortable" bandwidth is one that is less than about 40 percent of the center frequency. For JTIDS, this means that if we want to frequency-hop over a 250 MHz band, our center frequency should not be below about 600 MHz.

2.4 SPECTRUM ALLOCATION LIMITATIONS

As is currently the case with many of our natural resources, RF spectrum is limited. In recognition of this, it is international and national frequency management policy to make multiple use of already existing allocations if such expanded use can be achieved on a fully compatible and non-interfering basis.

Figure 1 shows the frequency allocation and typical utilization factor for the C-N-I systems currently in widespread use.

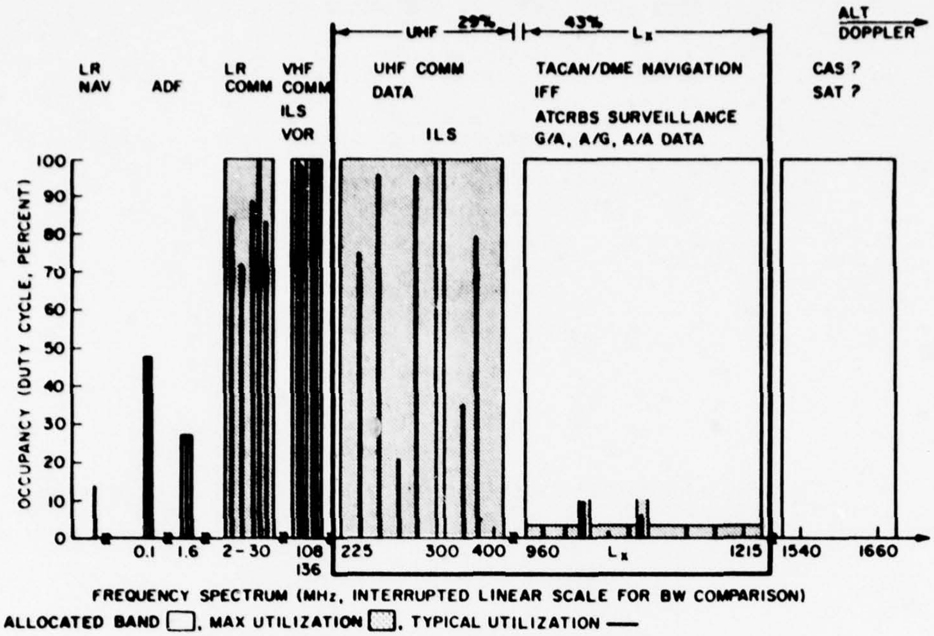


Figure 1. CNI Frequency Allocations

The chart of Figure 1 has been drawn with an interrupted linear scale for frequency so that the bandwidth for each of the bands can be readily compared while the ordinate shows the percentage utilization of each of the bands in two ways. The cross-hatched area shows that utilization which would exist if all allocated functions were operated simultaneously in a given area. The superimposed solid lines show the utilization in a typical operational area. We note that a large amount of spectrum in many choice frequency ranges has been allocated to aeronautical radio service, including communications, navigation, identification, and related functions. The figure shows the more attractive candidate bands. Of these, three bands are dominant; namely, the 225 to 400 MHz UHF communications band; the 960 to 1215 MHz Lx-Band, which provides Tacan/DME and IFF/ATCRBS functions; and the 1540 to 1660 MHz band, a part of which has been tentatively allocated to the ATA collision avoidance system and which has also been identified for L-Band satellite functions.

We note that over 70 percent of the bands depicted below 1660 MHz are currently allocated to C-N-I related functions, with over 40 percent dedicated to Tacan/DME and IFF/ATCRBS.

Taking into consideration the foregoing points, it is clear that we would like to operate at as low a frequency as possible in the 200 to 5000 MHz band, but with a center frequency above 600 MHz. In nearly all NATO countries, this latter frequency spectrum (600 to 900 MHz) is devoted to television. As we go up in frequency, the logical choice becomes the 960 to 1215 MHz aeronautical band. The next highest band, around 1500 MHz, would require about 1 1/2 - 2 times as much power, other things being equal.

One might be tempted to consider the 225 to 400 MHz band, decreasing the degree of frequency-hopping, but gaining a useful power advantage. However, the nature of the service density already in that band makes the addition of new services very difficult.

When we combine the considerations of limited spectrum the favorable electromagnetic characteristics, and the generally low frequency-time utilization factors, the 960 to 1215 MHz band is very attractive. It offers ample bandwidth (255 MHz), is lightly used (typical band duty factor is less than 0.1 percent), and further, involves large numbers of subscribers among broad classes of users. Each of these existing users bring an airborne hardware (volume, weight, antenna space, control and indicator space, etc.) suit whose functions can be partially (or in some instances fully) provided by the compatible CNI system. This is realized through the time shared utilization of receivers, transmitters, and signal/data processing hardware, thereby achieving an overall reduction of avionics. Similarly, each of these existing users represents an investment in equipment and operational experience which can stimulate strong support to a compatible CNI system, as compared to a revolutioaary non-compatible system.

Table 1 summarizes the key Lx-Band characteristics as they relate to the points covered above and to the ICNI/JTIDS requirements in general.

3. THE QUESTION OF COMPATIBILITY

When the 960 to 1215 MHz band was opened up to the civil aviation after World War II to provide DME service, it was generally recognized that this relatively vast spectrum allowed

Table 1. Lx-Band Characteristics

- INTERNATIONALLY ALLOCATED TO CNI FUNCTIONS
- AMPLE BANDWIDTH-255 MHZ
- IMPORTANT EXISTING FUNCTIONS
  - TACAN/DME & ASSOCIATED DATA LINK
  - IFF } INCLUDING TELEMETERED DATA
  - ATCRBS } INCLUDING TELEMETERED DATA
- BROAD CLASSES OF EXISTING USERS
  - TRI-SERVICE
  - AIRLINES
  - GENERAL AVIATION
- LOW UTILIZATION
- FAVORABLE PROPAGATION CHARACTERISTICS
- FAVORABLE HARDWARE CHARACTERISTICS
- IDEAL FOR GROWTH

many additional services to be implemented. RTCA SC-40 committee which met throughout 1948 to set DME standards rightly decided that they would not close the door to the addition of features such as bearing, ILS, data link, pictorial displays and other navigation services. One result of this was the requirement that receivers be able to tolerate large amounts of improperly coded (pulse-to-pulse spacing) pulse interference--later defined as 6000 single pulses, (with random pulse to pulse spacing) 60 dB louder than those with the desired code. This assumed that new services, to be added later, would use new pulse codes, and therefore be non-interfering.

As further evidence of this philosophy, the first Tacan sets built around 1951, had provision for an ILS operating in the 960 to 1215 MHz band; and in the mid-1950's the U.S. Navy developed a Tacan-compatible data link. Neither of these in any way interfered with the basic Tacan operation, but were abandoned for other reasons.

Thus, we can see that the concept of adding other services to the 960 to 1215 MHz band is not particularly new. The question that must now be addressed is "how many new pulses, of what signal format, can we tolerate?"

3.1 CNI/JTIDS - Lx-BAND COMPATIBILITY/SUSCEPTIBILITY

The existing CNI functions in the 960-1215 MHz Band are Tacan/DME and IFF/ATCRBS (AIMS). The Tacan/DME facilities provide navigation and data link functions on private duplex channels, while AIMS provides identification, surveillance, and telemetered data functions on a common duplex channel. In general, Tacan/DME services are excluded from the AIMS common channel frequencies so that the 255 MHz Lx-Band can be regarded as being composed of two subbands; namely, a subband of 230 MHz for private channels and a 20 MHz subband for the common channel. (The remaining 5 MHz is used initially for guard frequencies at the band extremes.)

To properly assess the impact of the new ICNI/JTIDS functions on conventional Tacan/DME and IFF/ATCRBS systems, the overall signal density must be calculated. This signal density will be a function of the new signal structure, maximum channel loadings, transmitted power levels/radiated spectrum, and geographical distributions of JTIDS and conventional equipments. This analysis is beyond the scope of this paper. We present some general comments, and briefly characterize the existing military and commercial equipments as to their tolerance to interference.

The nature of the new ICNI/JTIDS signal can best be characterized as; low duty cycle (≈ 10 percent), frequency hopped; wide band (≈ 3 MHz) spread spectrum. Radiated power levels vary from several hundred watts to over one kilowatt. Spectrum conserving Continuous Phase Shift Modulation (CPSM) is used to further contain the radiated spectrum. Although the basic carrier frequency is pseudo-randomly hopped across the 255 MHz band, guard bands of approximately ±20 MHz are placed around the 1030 and 1090 MHz IFF/ATCRB frequencies. Band edges are also protected with approximately ±9 MHz guard bands. This leaves us with approximately 150 MHz of usable spectrum across which the signal is actually hopped. Even at maximum loading conditions, the typical loading on any one-MHz channel is approximately 1500 pps.

The following paragraphs will, therefore, discuss the interference tolerance of conventional equipment in the more familiar terms of pulse interference.

3.2 AIRBORNE EQUIPMENTS

Military airborne Tacan units including the AN/ARN-21, AN/ARN-52, AN/ARN-90, AN/ARN-96, AN/ARN-84 and AN/ARN-118 are required to maintain specified performance in the presence of randomly spaced Tacan-type pulses which have a rate of 6,000 pps and an amplitude 60 dB above the desired signals.

ARINC Characteristic 568 (which governs civil DME equipments) requires specified performance (a 2-dB sensitivity reduction is allowed) in the presence of 7,200 Tacan type pulses 40 dB above the desired signal.

The highest estimated JTIDS on-channels pulse density (≈ 1500 pps) is well below these known airborne equipment tolerances. Further, it should be recognized that the 60 dB and 40 dB power ratios are equivalent to 1000:1 and 100:1 range disadvantage ratios, respectively. This would be equivalent to a Tacan interrogator trying to measure bearing and distance to a beacon 100 miles away in spite of the presence of JTIDS units within one mile. While this situation is unrealistically severe, the interrogator's measurement of bearing and distance would be unimpaired.

3.3 SURFACE EQUIPMENTS

No specific requirements in the specifications govern the operation of the surface equipments in the presence of random single pulse interference. From a review of the general design of the surface equipment we note that the effects of pulse interference are two-fold. One is caused by the effect of the interfering signals on the AGC circuitry, the other by the effect of the interfering signal on the echo rejection circuitry.

As the interference signal strength is increased further, the addition of the echo suppression circuits becomes evident. These circuits cause a desensitization of the receiver following reception of a high-level, noise pulse. The desensitization is proportional both in magnitude and duration to the amplitude of the noise pulse and acts to effectively reduce the receiver noise level for several tens of microseconds following each interference pulse. As this becomes more pronounced, the AGC compensates in the other direction, increasing the receiver sensitivity to maintain the constant decode rate. Tests have shown that interference pulse rates of 7200 pulses per second cause only a 1 dB variation in receiver sensitivity.

There is another effect, however. The blanking of the receiver immediately following reception of an interference pulse also causes a countdown of desired interrogation signals. Again, the effect is small and tolerable. For example, if a beacon is operated in the presence of interference sources within ten miles which generate 7200 pulses per second, an interrogator at 100 miles will be penalized with only 7 percent reply countdown interrogators are required by specification to tolerate 20 to 30 percent interference-caused countdown without performance degradation. The countdown caused by the much milder JTIDS interference would be minimal.

In addition to the analytically based performance capabilities of Tacan/DME equipments to JTIDS type pulse interference, extensive laboratory and flight testing has been completed to further validate the expected performance.

3.4 IFF/ACTRBS COMPATIBILITY

As indicated above, the new ICNI/JTIDS functions are introduced in only a 150 MHz portion of the band. These signals are intentionally excluded from the frequencies of the existing common IFF/ATCRBS channel (1030 ±5 MHz and 1090 ±5 MHz). This was done since the addition of this subband would not significantly increase the total spectrum available to these Lx functions, but more importantly, because the common channel functions of identification and surveillance are presently faced with problems of signal saturation. This is partly due to present system design (transponder systems develop "fruit" as a function of the square of the number of participants) and partly due to some misuse of the system. The JTIDS functions which use this common channel subband are, therefore, limited to and compatible with the identification surveillance functions which now exist in the band. No significant increase in interference is generated.

3.5 COMPATIBILITY TESTING

In addition to analytically determined performance capabilities of the airborne and surface Tacan/DME and IFF/ATCRBS equipments, extensive laboratory testing has been completed to further validate the expected performance.

In 1974 under the U.S. Navy's Integrated Tactical Air Control System (ITACS), ITT Avionics developed demonstration C-N-I terminals that utilized a spread-spectrum, frequency hopping signal structure operating in the 960 to 1215 MHz band. Compatibility testing on both airborne and surface Tacan/DME equipments and airborne IFF transponders was conducted. The results indicated that negligible levels of interference were caused to the equipments under test by the high levels of random pulse interference from the two ITACS terminals even when operating at high data rates (approximately 60 kbps), and high power levels.

In 1975, the FAA conducted laboratory tests of a precision DME set operating in the Lx-Band, intended for use with MLS. The system used phase-coded pulses very similar to those of JTIDS, but without frequency hopping. No mutual interference effects were found.

More recently, as flight worthy JTIDS terminals became available, extensive airborne testing was conducted. During these tests, FAA certification aircraft (used to checkout airborne and surface Tacan/DME systems), were used as victim equipments. The basic testing involved the generation of JTIDS pulse traffic at various levels of output power and signal density. This simulated range differences (power level variations) and data rate differences (signal density variations) that might typically be encountered by conventionally equipped platforms operating in a JTIDS area.

These tests were conducted under carefully controlled conditions and involved personnel from the FAA, OTP and the various military services. The results validate the expected performance and show that the JTIDS signal traffic will have no deleterious effect on existing services.

3.6 JTIDS SUSCEPTIBILITY

One issue of spectrum sharing is to determine the effects of existing Tacan/DME and IFF/ATCRBS signals on JTIDS operation.

3.6.1 Tacan/DME Interference Effects

JTIDS operates with high levels of processing gain afforded by advanced signal processing techniques. Even for those cases where the high-power ground station (5 to 10 kW) acts as a close-in interfering source to the desired far-out signal, the JTIDS architecture has very powerful errata correcting capabilities afforded by its Reed-Solomon error correcting code. Specifically, over 50 percent of the pulses of any message block can be lost and the message fully recovered by the code.

In summary, we see that the JTIDS signal is not susceptible to Tacan/DME interference for all practical cases of interest.

3.6.2 IFF/ATCRBS Interference Effects

As indicated earlier, the JTIDS system excludes the IFF/ATCRBS frequency bands. This will, therefore, preclude any levels of pulse interference from reducing JTIDS performance.

4. CONCLUSIONS

JTIDS/Integrated CNI operates in the 960 to 1215 MHz portion of the electromagnetic spectrum. We have briefly reviewed the general and specific characteristics of candidate frequency bands and shown that the Lx-960 to 1215 MHz portion is well suited to ICNI/JTIDS application. The compatibility/susceptibility of the JTIDS versus Tacan/DME and IFF/ATCRBS is also shown to be acceptable.

This choice portion of spectrum is already internationally allocated to navigation functions. Within some broad guidelines, of pseudo-radom frequency hopped low duty cycle signal structures, it can be effectively time shared by the JTIDS system and thereby afford the benefits of both integrated CNI and advanced high performance JTIDS applications.

A question may remain as to whether, if JTIDS is fully compatible with DME/Tacan, JTIDS is the proper system to help "fill up" this band, perhaps prelcuding the later addition of still other systems. The best answer to this, we believe, is that in the 20 years since international agreement was reached on the use of the 900 to 1215 MHz band no other system has been seriously considered.

JTIDS SIGNAL STRUCTURE

DR. GERALD J. PALATUCCI

COMMUNICATIONS, COMMAND & CONTROL DIVISION
COMMUNICATION NAVIGATION TECHNOLOGY DIRECTORATE
U.S. NAVAL AIR DEVELOPMENT CENTER
WARMINSTER, PA 18974

I. Introduction

The JTIDS signal structure is the product of many engineering hands in the military organizations as well as the industrial organizations. The JTIDS signal structure represents the optimization of a waveform design against many complex and interrelated constraints. The ultimate success of the system will be measured within the economic constraint of performance against threat for the given resources invested.

This signal structure section of the paper will describe in general terms a hybrid frequency hopping, time hopping, and direct sequence pseudo noise (PN) spread spectrum system for communication, navigation, and identification applications.

These applications have often been described in terms of ICNI (Integrated Communication, Navigation and Identification) systems. A few words of orientation are provided here to distinguish between two active ICNI concepts which directly relate to the design constraints. The first ICNI concept is referred to as "waveform ICNI" and relates to the design of any advanced signalling structure which may be used to satisfy data transmission needs, (C) time of arrival measurement needs (N) and fine synchronization needs (I) in one integrated design. The second ICNI concept is referred to as the "hardware ICNI" concept and has to do with a given set of radio transceiver/modem hardware processing more than one function. For example, a set of terminal hardware used to process both JTIDS and TACAN signals in the Lx (960 MHz to 1215 MHz) band. This hardware ICNI concept is an extremely important one which is needed in the interim era between older system phase out and new system introduction. It represents a politico - engineering solution to this always present problem.

This section is divided into three major parts. The first part deals with the primary design constraints placed on the system at the outset of the waveform design. These constraints have been classified as either fundamental bounds or requirements bounds. The fundamental bounds are ones which occur because of technological state-of-the-art or because of resource bounds such as dollar cost or frequency spectrum. The requirement bounds occur because of specifications placed on the performance of the system. The second part deals with some primary characteristics of the final signal structure product such as channel coding, data modulation and RF modulation. The third and final part of this section deals with other important signal structure concerns for JTIDS.

II. Design Constraints

Fundamental Bounds

The bandwidth available for implementation of the signal structure was chosen for reasons which are given elsewhere. Under benign operating conditions certain parts of this band are excluded from transmission and reception to minimize the possibility of interference with ATCRBS (Air Traffic Control Radar Beacon System 1030 MHz and 1090 MHz) and to minimize band edge spillover. The effective spread spectrum bandwidth in this case is about 200 MHz. This large bandwidth is able to be included in the JTIDS signal structure by utilizing a number of spread spectrum techniques. It is unlikely that a direct sequence system could be implemented over this whole frequency range using active correlators which would meet the requirements of JTIDS. A hybrid system, on the other hand, which combines frequency hopping techniques as well as direct sequence techniques has advantages. The noncoherent frequency hopping found in JTIDS is the first level of spectrum spreading. This frequency hopping provides the advantages of relatively rapid synchronization and nominal multipath immunity. The practical consequence of this approach results in a pulsed system rather than a CW one.

The technology available to do passive correlation drives the time bandwidth (TB) product to be used in the pulse compression techniques. From the cost and practicality point of view it is always good to have at least one mature technology available to do pulse compression. In the case of JTIDS either LSI digital correlators or surface acoustic wave (SAW) devices may be used to do the pulse compression. Two other technologies may soon be applicable - charge coupled devices and SAW convolvers.

The third level of spectrum spreading employed in JTIDS is referred to as time hopping. This represents a technique which is used to ensure that occurrences of various JTIDS pulses fall in a time pattern which is seen as random and unpredictable by a non-member of the system.

The total cost of the JTIDS hardware is minimized by employing the hardware ICNI concept. In order to share the RF power amplifier and receiver hardware certain conventional function modes must be retained in the JTIDS hardware. For example, the pulse envelope shaping for TACAN and the 1 MHz channelization for TACAN reception. The duty factor of the JTIDS system with either singlet or doublet pulses (to be discussed below) is higher than for the other Lx band pulsed systems TACAN and IFF. Hence, the JTIDS transmitter may be shared with these other functions, in principle, at no average power duty burden on the transmitter design.

In a benign environment, the TACAN and ATCRBS RF activity represents a pulse noise background in which JTIDS must operate. This area of concern for the JTIDS signal structure design has been referred to as "susceptibility" and has to do with the tolerance of JTIDS to the in band "friendly jammer". Suffice it to say that the JTIDS frequency hopping and data encoding prevent the interference from effecting the performance described in the JTIDS specification.

On the other side of the frequency sharing coin is the concern over degrading TACAN, ATCRBS or IFF services. This area provided one of the strongest design bounds for JTIDS and has been referred to as JTIDS compatability. A long history of experimental results with various waveform pulse candidates has resulted in the molding of many JTIDS parameters. The pulse width for JTIDS has been arrived at for a number of reasons. One has to do with the TACAN pulse pair decoding circuitry. Certain pulse widths were taboo for JTIDS because of this possible interference mode with TACAN. Another reason for the pulse width of JTIDS has to do with the pulse instantaneous bandwidth and the ability to measure time of arrival (TOA) for navigation applications and achieve some multipath immunity. The number of PN chips in the pulse which is tied to the correlator time-bandwidth (TB) product is the important thread which ties all the pulse parameters together. Finally, the RF modulation for the PN chips of the pulse was chosen to be continuous-phase shift modulation (CPSM). This mode of quadriphase modulation is sometimes referred to as minimum shift keying (MSK) and is noted for $1/f^4$ frequency spectral rolloff. This characteristic of the JTIDS signal structure mitigates against in-band and out-of-band interference.

Required Performance Bounds

A message error rate (MER) and a false message acceptance rate (FMAR) were specified as requirements for the JTIDS signal structure. Furthermore, the specified levels of message quality were to be guaranteed in specified interference environments. This anti-jamming margin or J/S (jammer power to signal power) is a quality measure which has been continuously employed (along with other factors) to evaluate the various predecessors and the present JTIDS waveform. It is important, however, to note that most quality measures do not completely describe a given system. In the case of JTIDS the final measure of system goodness is the measure of the relative total economic burden on an adversary to defeat the system capability.

The quality measures for message reception (MER and FMAR) are used to determine the required Eb/No (energy per bit per noise hertz bandwidth). The probability that a message is received correctly is dependent on the probability that synchronization is acquired and the probability that data within the message is received correctly given that synchronization has taken place. Evaluation of this conditional probability equation along with one that specifies the false message acceptance rate is routinely done to determine the behavior of the system to various jamming strategies and to various other factors such as implementation losses or receiver sensitivity. Generally, the probabilities are budgeted between reading synchronization and reading data. One budget would allow the probability of synchronization to decrease and the probability of reading data to increase while keeping the product constant. Originally, the budget was at the discretion of the JTIDS system designers and has been chosen to minimize hardware costs.

Stringent required bounds were placed on the JTIDS signal structure to accommodate the needs of navigation applications. It was necessary to measure time of arrival very accurately to guarantee successful weapon delivery. The JTIDS TOA fixes of own position against others or against beacons are the inputs to navigation algorithms which allow the accurate computation of relative navigation coordinates or of accurate geodetic coordinates especially when used in conjunction with satellite navigation systems. If the TOA accuracy was not adequate these errors could propagate in the computations to result in unacceptable position coordinates. The JTIDS signal structure allows a refinement to conventional navigation accuracy of an order of magnitude.

The final performance requirement to be discussed here is the time specified to acquire synchronization of the system. The JTIDS signal structure has been designed to statistically guarantee acquisition in a very short time. The authorized entering JTIDS community member is able to program his JTIDS hardware to receive a sequence of pulses whose energy may be summed before detection. In one mode, accurate RTT (round trip timing) messages may be exchanged with a synchronized member or, in another mode, navigation information may be used for passive entry.

III. JTIDS Waveform Characteristics

Frequency/Time/Autocorrelation Domain Properties

The power spectral density of the composite JTIDS signal structure (or the JTIDS signature) is optimized in the sense that it has been made as uniform as practical. The PN spreading, frequency hopping, and time hopping cause the autocorrelation function of the JTIDS signal to be substantially impulse-like and hence results in a flat power spectral density. If the power spectral density of a spread spectrum system is "colored" in any way, it represents a potential strategy for an adversary to undertake. The jamming strategy would include matching to this colored power spectral density. This subject is further discussed under the "optimum jammer" section below.

Channel Coding and Data Modulation

The JTIDS signal structure employs concatenated encoding to satisfy the many design constraints mentioned in Part II. The inner most coding is that of a 32' ary alphabet implemented with cyclic code shift keying (CCSK). A special code called the data or So code is arranged in any one of 32 phase positions to designate the 32 members of the alphabet. The task at the detector is to decide which character was sent. The detector may be a maximum likelihood one in which 32 comparisons are made or the detector could be implemented simply with threshold detection circuitry. A double length SAW correlator may be used as a data decoder, for example.

The next level of data encoding has to do with implementing a long block code used for residual data error detection capability.

The next higher level of encoding is the use of a Reed Solomon 32' ary block code. The Reed Solomon code capability is directly related to the use of a 32' ary alphabet in the channel. The Reed Solomon code in JTIDS is a character correcting (vs. bit error correcting) code whose performance is achieved when matched to a channel which produces the corresponding character errors and/or character erasures. (The erasure mode or no decision mode at the detector yields substantial performance improvement over decoding without erasures.)

The Reed Solomon code used in JTIDS is a BCH (Bose Chaudhuri Hocquenghen) code which, for a given code rate (e.g., 1/2), alphabet size (e.g., 32' ary) and block size (e.g., 15 information characters) represents a perfect code. That is, there can be no better performing block code. The Reed Solomon encoder adds code characters to the information characters and protects the information in a systematic block code. Thus, if desired, the information characters can be directly recovered without the use of a decoder. In a benign environment an austere or special mode (e.g., degraded) user can receive and read Reed Solomon encoded messages without decoding. Concealment of information by encryption is a separate and optional operation.

The Reed Solomon code provides maximum error correction capability. An error occurs when one character is transmitted and the receiver erroneously decides that another character was received. In the JTIDS code, 15 information characters are protected with 16 parity characters for a total block size of 31 characters. Practical operations are not confined to cases with all errors or all erasures but with combinations of errors and erasures collectively called errata. As long as the sum of any combination of erasures and twice the errors is less than or equal to 16 the correct information will be recovered.

Erasures occur for the following types of events. Transmitter erasures occur when character reception is prevented due to coincident transmitter triggered blanking. In this case, no reception occurs, no decision is possible and declaration of erasure is the proper action. Receiver erasures, on the other hand, occur when character reception is prevented due to other coincident receptions on other frequencies which occupy available receiver channels. Again, no reception occurs, no decision is possible, and declaration of erasures is the proper action. The final erasure type is the interference type and occurs when RF interference significantly reduces the probability of correct decisions. Since the JTIDS Reed Solomon decoder can correct twice as many erasures as errors there is value in providing erasure detection capability.

The described erasure correction capability and associated 32' ary modulation are critical to multi-net operations. The capacity of a multi-net system is based on both the level of mutual interference generated and the ability to tolerate such interference. For any scenario of geographic distribution of data sources and data rates, all types of mutual interference are inversely proportional to the number of bits per pulse. It is fact that the total JTIDS mutual interference effects represent only a fraction of the total interference budget.

The Reed Solomon (31,15) code coupled with the 32' ary data modulation are important for TACAN mutual compatibility, related system capacity and TOA accuracy. For any given data rate and pulse dimensions, JTIDS interference to TACAN receivers is inversely related to the number of bits per pulse. Similarly, for any given pulse dimensions the Reed Solomon encoder/decoder can tolerate the interference from TACAN transmitters. If interference tolerance is limited, the system bandwidth must be limited. However, the ability of a signal structure to produce accurate TOA measurements (especially in a multipath environment) is directly related to pulse bandwidth. Thus, all other things being equal, the higher tolerance of Reed Solomon coding allows a system to employ wider bandwidth pulses in a TACAN or pulse interference environment and produce correspondingly higher TOA accuracies for navigation, relative navigation, and/or platform control. Furthermore, the Reed Solomon code is not sensitive to burst errors but only to the total errata in a block.

Certain JTIDS operations involve the recognition of short addressed interrogation messages and the immediate generation of a reply so that the total round trip operation is completed within one time slot. This is accomplished by including along with the main block decoder algorithms, a simple block recoginizer algorithm which is channel programmed to recognize the specific short message composed of own address and appropriate interrogation label.

One further encoding scheme is used in JTIDS to ensure that time domain characteristics of the signal structure are optimized against burst jamming strategies. Interleaving is employed to scramble the time channel position of characters of multiple code blocks in longer messages just before transmitting and interleaving is employed to descramble just after buffered message reception.

IV. Other Concerns Addressed by the JTIDS Signal Structure

The "optimum jammer" concept is one which recognizes the ability of an adversary to obtain system parameters through dedicated measurements. The jamming margin against this dedicated jammer is computed by conceptually allowing the jammer all knowledge of the spread spectrum design parameters except for parameters assigned cryptographically to the system on a periodic basis.

JTIDS: HARDWARE DEVELOPMENT

David A. Herrelko, Captain, USAF
JTIDS Joint Program Office
Electronic Systems Division
Hanscom Air Force Base, Massachusetts

The integrated communication, navigation, and identification features of JTIDS will serve as a force multiplier to the tactical arms of the Army, Navy, Marine Corps, and Air Force. The expected requirements of the Services and of allied nations will include installation on fighter, cargo, and command/control aircraft; on ships; on tracked vehicles; and at fixed bases. Installations will occur in many cases on a retrofit basis, and must disturb the existing weapon systems and C^3 systems as little as possible.

Given the evolution of JTIDS described by Mr. Vaughn earlier in this AGARDograph, and given the array of installations and missions for which JTIDS is a candidate, it is prudent to employ a sequenced development program. The classical unimodal diagram of a program life cycle cannot apply here; rather, JTIDS must evolve in an overlapping continuum of developments. An ordered cascade of projects--each of which is keyed to a technology, a platform, a mission application, or an enhancement--must be undertaken. Throughout these multiple developments, the JTIDS procurement policy is to encourage competition, to "try before fly", and to "fly before buy".

JTIDS will be developed in two major phases. Phase 1 will include full-scale development of the basic information distribution system, with added capabilities for relative navigation, and digital voice capability. It is designed to provide hardware for implementation in the early 1980's. As pointed out in Mr. Eisenberg's system overview, certain limitations are inherent in Phase I, and these will be addressed in Phase II. The research, development, and experimentation of Phase II is intended to alleviate many of the limitations, and to incorporate other capabilities useful in a tactical environment.

Beyond actual terminal development, the JTIDS JPO is pursuing a number of technical investigations that, while not central to precision navigation, are of great importance in exploiting fully the powers of JTIDS. Among these interest areas are digital voice, error detection and correction, data encryption, surface acoustic wave devices, electron bombarded semiconductors, power amplifier modules, and adaptive nulling antennas.

In the papers that follow, major JTIDS contractors discuss their approaches and achievements in hardware development aimed at the Phase I and Phase II requirements.

COMMAND AND CONTROL TERMINALS

Ara N Boyajian
Hughes Aircraft Company
PO Box 3310
Fullerton, CA 92634

JTIDS Terminal developments at Hughes Aircraft Company have progressed through three equipment configurations commencing with the development of "A" waveform TDMA Terminals for the E-3A Airborne Warning and Control system. The second major Terminal Development effort was the conversion of the "A" waveform Terminals to the "B" waveform configuration which was adopted as the JTIDS standard signal structure subsequent to the TDMA contract award. Additional development efforts incorporating functional improvements and the use of more advanced technology resulted in third generation equipments, the Hughes Improved Terminal (HIT). The following discussion briefly describes these configurations and this progression in the maturing process of the JTIDS through January 1978.

TDMA for E-3A (Radio Set AN/ARC-181)

Hughes Aircraft Company was contracted in September 1974 to develop Command and Control Class I TDMA Terminals for integration into the E-3A AWACS Aircraft. The award was the result of a competition for pre-production equipments based on a minimum cost approach utilizing state-of-the-art technology with size and weight considerations having secondary importance. Component development efforts were constrained to only the high power transistors.

Laboratory testing at Hughes Aircraft Company which preceded equipment deliveries consisted of performance tests for engineering design verification and acceptance, pre-qualification safety of flight tests, and compatibility tests with Air Traffic Control equipment (airborne TACAN/DME, ground beacons, and IFF interrogators/transponders). Terminal acceptance tests verified terminal functional operation, operator interfaces, initialization and transmitter operation (power output, frequency and spectrum). Network level testing verified the performance of the terminal operating in a two-terminal network. Specific tests included:

- Net entry and synchronization in all modes
- Message processing and through-put
- Sensitivity (Minimum Discernable Signal)
- Receiver dynamic range
- Anti-jam margin

All primary system performance requirements, pre-qualification tests and compatibility tests were successfully verified.

Initial equipment deliveries were made to the Boeing Aerospace Company in January 1977 for integration into their Development Integration Test Facility at Seattle for software verification and operability testing. Additional terminals were delivered to Boeing in June 1977 for installation on the E-3A test system aircraft and a government furnished NKC-135 aircraft for ground radiation and flight testing.

An extensive flight test program was successfully concluded in September 1977. These tests conducted by the Boeing Aerospace Company demonstrated the operability of TDMA in the E-3A in air-to-air, air-to-ground, and air-to-air-to-ground (relay) scenarios. The flight test program consisted of twenty-four flights demonstrating TDMA performance. The measured system performance indicated that there is sufficient link margin to ensure that the specified performance is attained in all modes of operation in the AWACS environment. All test objectives and performance criteria were met or exceeded without any system failures in flight with over 200 hours of airborne operation. The observed reliability of the TDMA system in over 7500 hours of operation at Boeing exceeded specification requirements. The following listing summarizes these flight tests.

- Three terminal network operation demonstrated
- Relay operations over 500 mile range demonstrated
- Ability to time shift a network demonstrated
- Network flexibility demonstrated – no advance or current knowledge about other users in network is required for entry
- Warm up time for master and user terminals for transmit and receive determined
- Double pulse signal structure vs single pulse performance evaluated
- Performance in turns evaluated in both hi and lo power modes
  - Demonstrated that link performance meets specification requirements
  - Limited engineering data obtained for greater angles of bank

At the completion of the Flight Test Program all of the "A" waveform terminals were returned to Hughes for modification to the "B" waveform configuration. The first two "B" waveform Terminals were delivered to Boeing in January 1978.

System Configuration

Figure I shows the units that comprise the AWACS TDMA configuration and the inter-relationship between the units. The system was developed for operation in either a high power or low power mode, radiating through either of two antennas or simultaneously through both.

The Digital Computer or Communications Processor is an IBM AP-1A general purpose computer which interfaces with the AWACS central computer. It controls the terminal operation and performs all terminal processing functions. It monitors and controls system self-test functions with an external data recording system, and provides for information exchange with the Radio Set Control.

The Radio Set Control which comprises the man/machine interface provides for operator terminal initialization, power control, and fault isolation.

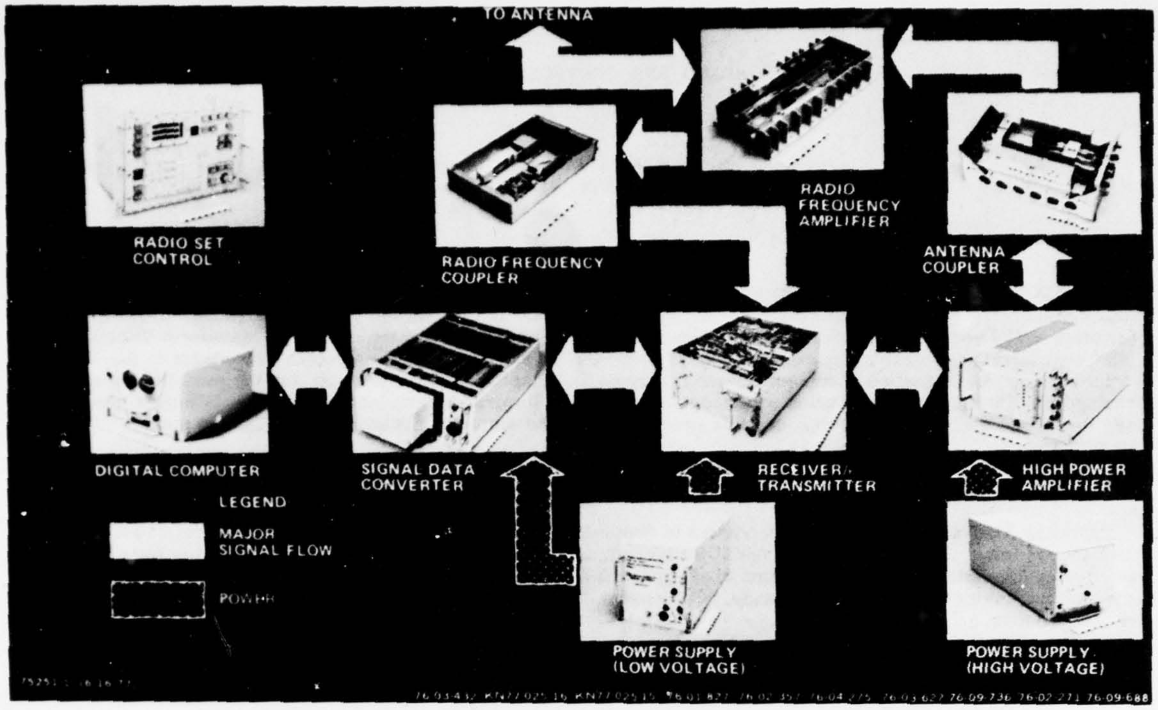


Figure 1. Radio Set AN/ARC-181

The Signal Data Converter interchanges messages with the Digital Computer, provides for error correction encoding and decoding, performs spectrum spreading on the messages, and modulates/demodulates the data. It also generates acquisition preambles for transmission and correlates received preambles, and houses and interfaces with a Secure Data Unit subassembly. The Secure Data Unit, designed and developed by Hughes Aircraft Company for the E-3A program, is utilized for all JTIDS developmental applications.

The Transmitter/Receiver accepts outgoing messages from the Signal Data Converter, up-converts them to L band and provides r.f. power amplification for output to the antenna(s). On the receiver side it receives incoming signals, demodulates the acquisition preamble, down converts the data portion of each message, and outputs it to the Signal Data Converter.

The Antenna Coupler provides r.f. coupling to one or two omnidirectional antennas. In addition, r.f. filtering for in-band and out of band emissions is provided.

The Low Voltage Power Supply accepts 3 phase 400 Hz prime power and converts it to the appropriate d.c. voltages for the terminal.

The High Power Amplifier provides a higher r.f. power output than is available from the TR. Associated with the High Power Amplifier is the High Voltage Power Supply which concerts 3 phase 400 Hz to the appropriate d.c. voltages.

The pre-amplifiers provided for E-3A include two each Radio Frequency Amplifiers which provide low noise pre-amplification of the receive signals and one each Radio Frequency Coupler which combines receive signals.

Differences between the "A" and "B" waveform terminal configuration are reflected in the Signal Data Converter and Receiver/Transmitter units. All other units are unmodified and are compatible with either configuration. Aside from the different waveform signal structure, the "B" waveform terminal is implemented with a Hughes developed Reed-Solomon encoding/decoding device for forward error correction utilizing a 32-ary code operating at a 5 MHz chip rate. The "A" waveform terminal utilized a Viterbi encoding/decoding device employing an 8-ary code operating at a 3 MHz chip rate. Both configurations were implemented for selectable operation with either the single-pulse or double-pulse signal structure with redundant information carried in the second pulse. Anti-Jam protection is provided through the incorporation of several techniques which apply to both the "A" and "B" waveform terminals.

Hughes Improved Terminal (HIT)

The Hughes Improved Terminal is an outgrowth of Company funded development efforts which resulted in a new generation reduced size JTIDS Terminal. The Hughes Improved Terminal retains all of the features and performance characteristics of the Class I Terminal used in the E-3A but with the utilization of superior technology which allowed for the physical compression. Hughes is currently under contract for delivery of HIT Terminals to ESD commencing in mid 1978. The first three terminals, Engineering Development Models, are currently in laboratory testing at Hughes undergoing system functional and performance tests. Follow-on HIT Terminals are being manufactured in Hughes production facilities with initial deliveries commencing in the Spring of 1979.

System Configuration

The Hughes Improved Terminal consisting of a Transceiver-Processor unit and a Low Power Amplifier/Power Supply is shown pictorially in Figure 2, and in the functional block diagram of Figure 3, along with the optional units

which comprise a full-up terminal for E-3A applications. The RF output/input of the terminal is at an Antenna Interface unit through RF cable connections to the system antennas. At the opposite end of the terminal, the inputs and outputs are through an Interface Adapter Unit, an I/O Multiplexer Bus, and the Unformatted Message Element. An operator interface is available through the Radio Set Control. Interfaces to a Recorder/Reproducer are also provided.

The Transceiver/Processor Unit is comprised of a Communications Processor, Signal Processor and Transmitter/Receiver which perform all of the functions of the Digital Computer, Signal Data Converter and Transmitter/Receiver of the E-3A TDMA Terminal (AN/ARC-181). These functions are integrated in a single unit taking advantage of miniaturization and eliminating the volume of two additional enclosures and interfacing cabling.

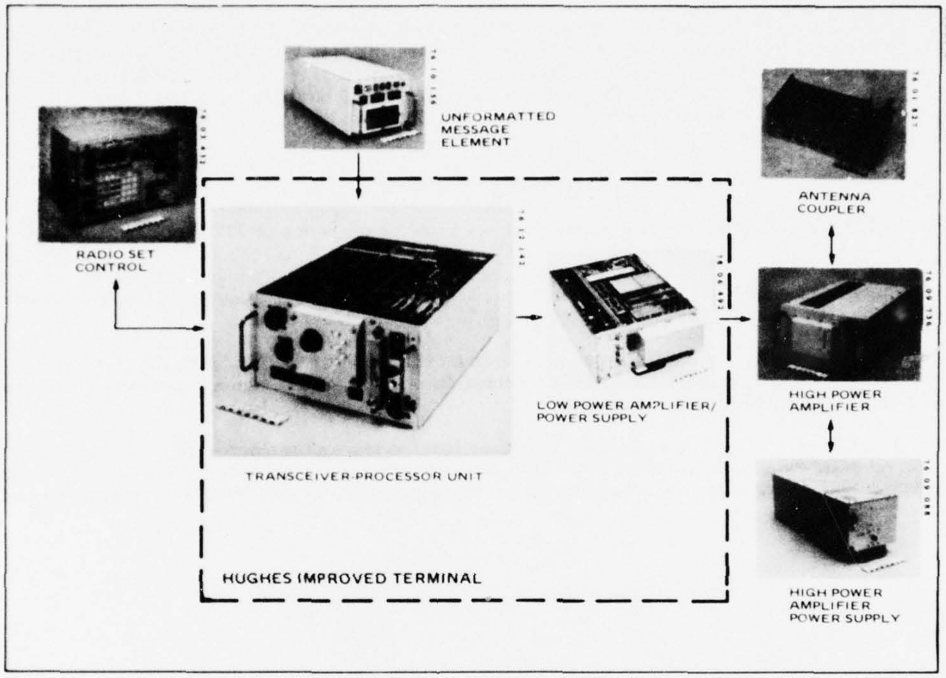


Figure 2. Hughes Improved Terminal and Optional Interfacing Elements

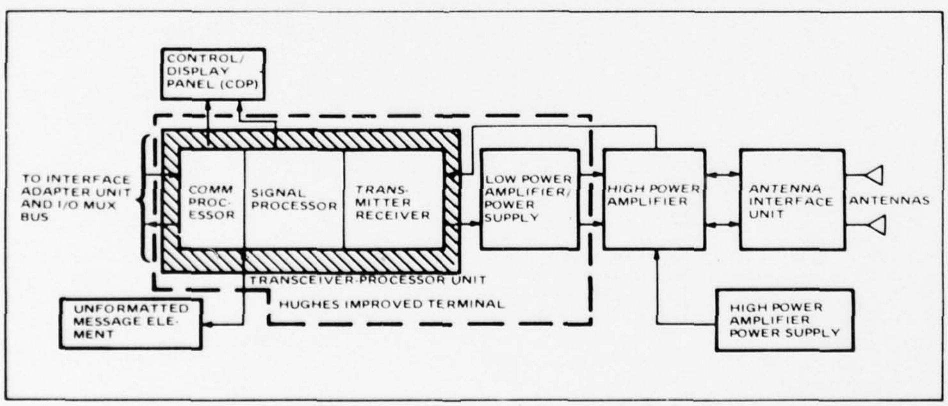


Figure 3. Hughes Improved Terminal Functional Block Diagram

The Communications Processor is an advanced version of the Hughes-developed HMP-1116 mini-computer which offers the advantages of an established architecture with high performance and well developed support software. The HMP-1116 Communications Processor is a 16-bit microprocessor with a 32-bit Arithmetic Logic Unit and a full set of double precision instructions in the repertoire. The microprogram memory capacity is 151 instructions. The HMP-1116 utilizes the newest LSI semiconductor devices with many performance features included as standard (floating point arithmetic, byte handling, list processing instructions, privileged operating modes, and memory parity).

The Low Power Amplifier and the Low Voltage Power Supply are contained in one enclosure which comprises the second unit of the HIT. The Low Voltage Power Supply converts the prime AC power into DC power for use by all three elements in the Transceiver-Processor Unit and the Low Power Amplifier.

The Hughes Improved Terminal is implemented for the waveform B signal structure, which is compatible with the basic TDMA network architecture and time-ordered structure. The B waveform structure in HIT is selectable as either the single-pulse type 2B or double-pulse type 3B. A Reed-Solomon encoder is used for forward error correction coding. In the HIT implementation, each data symbol represents 5 bits in the form of a byte. A 32-ary modulation is employed, wherein 32 distinct symbols of 32 bits each comprise the alphabet employed by the terminal. The 32-ary symbols are transmitted at a rate of 5 megabits per second.

Technology Improvements

The following areas of technology improvements are cited as examples of JTIDS related developments.

Error Detection and Correction – The terminal uses Reed-Solomon error detection and correction, which has previously required significantly large computer capability to perform in quasi-real time. A hardware encoder-decoder using catalog LSI has been developed, and requires only 4 (5-1/2" x 5-1/2") cards.

Surface Acoustic Wave Devices – S.A.W. filters have been fabricated at 315 MHz which provide the rapid roll off required to achieve the necessary JTIDS spectral characteristics. These single component devices replace the extensive filtering required by conventional techniques.

L-Band Power Amplifiers – Since jam resistance is proportional to transmitter power, the maximum practical output is desired. An all solid amplifier generates more than 1600 watts (at 20% duty) over a 250 MHz bandwidth. This exceeds the bandwidth achievable with gridded tubes, and is smaller than L-Band TWT's previously needed to obtain such power.

TACTICAL TERMINALS

Ara N Boyajian
Hughes Aircraft Company
PO Box 3310
Fullerton, CA 92934

The Hughes Improved Terminal (HIT), described previously in Section IVD(1)a, provides high anti-jam and processing capability in a Terminal applicable to Tactical applications. The HIT was the result of Hughes development efforts to reduce the physical size of the larger E-3A (AWACS) Command and Control Terminals while retaining its full performance characteristics. As previously mentioned, the first three Engineering Development Model Terminals are currently in laboratory testing at Hughes undergoing system functional and performance tests. The follow-on Hughes Improved Terminals are being manufactured in Hughes production facilities with initial deliveries commencing in the spring of 1979.

The utilization of the HIT in either a Tactical or Command and Control application is determined by the RF power levels that are radiated and the degree of man-machine interaction required. The Transceiver-Processor Unit of the HIT with its Low Power Amplifier/Power Supply Unit and appropriate controls and displays represents a Tactical Terminal configuration (Figure 1). The addition of a High Power Amplifier and Power Supply; and integration with Command and Control Displays and Controls converts the HIT for Command applications. If additional Navigation and Identification functions such as TACAN, IFF or GPS terminal functions are integrated within the Tactical Terminals, the differences between Class I and Class II Terminals become more pronounced.

The HIT has been considered for a Tactical flight demonstration program on the F-15A aircraft and for certain other test programs where the physical constraints are not severe. It is anticipated that a derivative of the Hughes Improved Terminal would be configured as a candidate for a Full Scale Development program. A Full Scale Develop-

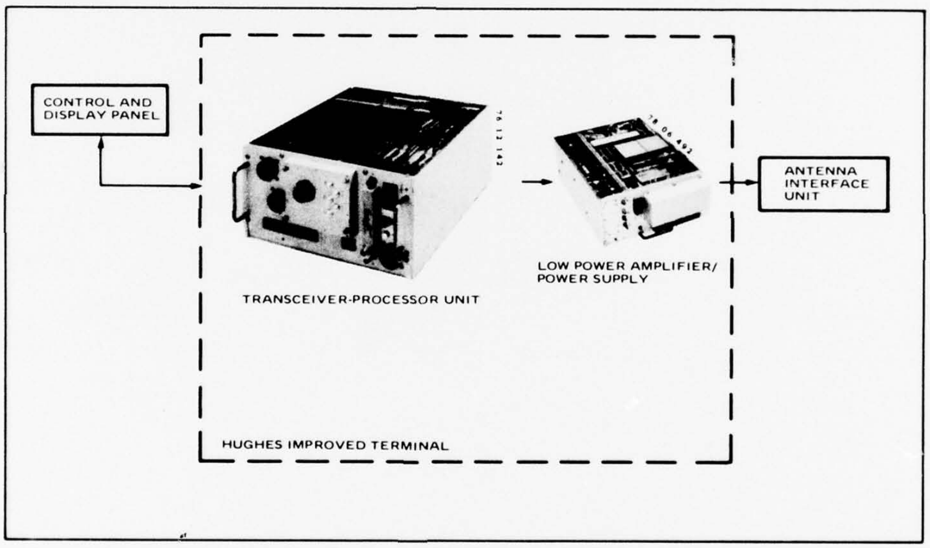


Figure 1. Hughes Improved Terminal, Tactical Configuration

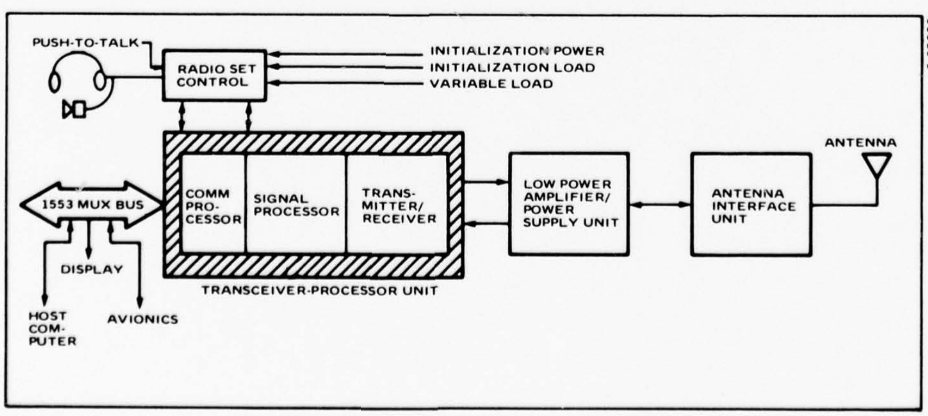


Figure 2. Functional Block Diagram, Baseline System

ment JTIDS Class 2 implementation would require further size reductions in the HIT along with the incorporation of additional functions and unique interfaces that would be specified. If specified, an expanded operational capability can be provided through modifications of the basic TDMA system to provide higher data capacity and flexible message formats. These modifications, which have been defined and evaluated, would provide Class 2 ATDMA capability with relatively low risk.

The block diagram of Figure 2 shows a baseline Terminal with a Radio Set Control and an interface with a MIL-STD-1553 multiplex bus that most of the advanced platform Avionics will utilize. For those platforms that do not have a 1553 interface, the flexibility exists for the inclusion of an appropriate interface module within the terminal, thus precluding the need for an additional interface unit.

AN/URQ-28 JTIDS CLASS 2 TACTICAL TERMINAL

R. AUGHEY - G. SOLOWEY

The Singer Company
Kearfott Division
150 Totowa Road
Wayne, New Jersey 07470

General

The JTIDS Tactical Aircraft Terminal AN/URQ-28 provides a low-detectability, cryptographically secure, anti-jam, integrated communication and relative navigation system within the Receiver-Transmitter housed in one 1 1/4 ATR case, suitable for military aircraft. The system can operate in a mix of airborne, fixed wing and rotary wing, sea based and ground based vehicles. The AN/URQ-28 provides a nodeless, time ordered (TDMA) two-way, high capacity digital data link for real time communication between all community members. The Terminal also provides high accuracy relative navigation in both active and passive modes. Automatic identification of all community members is provided by predetermined specific time slot assignments. The AN/URQ-28 has been designed in accordance with JTIDS 2B Waveform signal structure specification. Some of the major performance characteristics of the AN/URQ-28 are presented in Table 1.

AN/URQ-28 Description

TACAN/TDMA Common Circuit Design

The evolutionary design concept of the AN/URQ-28 is exemplified by provisions for the substitution of the terminal in place of an AN-ARN-52 or AN/ARN-84 TACAN set in a military aircraft thereby keeping the full TACAN capability and adding the communication and relative navigation functions of the AN/URQ-28.

In order to provide these multiple functions in a 1 1/4 ATR case Singer-Kearfott has achieved a unique, Kalman filter hardware/software design approach to the TACAN and TDMA relative navigation functions with significantly reduced hardware complexity.

Additionally, extensive use of the following common hardware circuitry for TACAN and TDMA function has been made to reduce weight and volume.

TACAN/TDMA Common Circuits

1. Antenna
2. Antenna Switch
3. Protective Limiter
4. Tunable Filters
5. RF Preamplifier
6. Mixers
7. Frequency Synthesizer
8. Computer/Signal Processor

The AN/URQ-28 terminal overall block diagram is shown in Figure 1. The terminal consists of a multi-function transceiver and signal processor plus an external RF spectrum control filter (notched at 1030 and 1090 mhz to prevent IFF interference) and a cockpit mounted Mode Control Unit (MCU). Also shown in Figure 1 is the Control Display Unit. This support equipment will be eliminated for those aircraft installations where an integrated control system already exists.

TABLE I

AN/URQ-28 PERFORMANCE CHARACTERISTICS

RF PEAK POWER OUTPUT:	800 or 120 Watts (Solid State)
FREQUENCY BAND:	960-1215 MHz (Notches at 1030 & 1090 MHz)
MODULATION:	CPSM (Controlled Phase Shift Modulation)
COMMUNICATION CAPACITY:	57,000 Bits/Sec
DATA MODULATION:	32 ARY
OPERATING RANGE:	300 NM LOS - 500 NM with Relay
CRYPTOGRAPHY:	NSA Compatible
TIME SLOT DURATION:	7.8125 msec
SYSTEM CYCLE:	1536 Slots (12 Seconds)
EPOCH:	64 Cycles
MESSAGE TYPES:	1. Formatted 2. Free Text Coded 3. Free Text Uncoded 4. RTT (Round Trip Timing)
ANTENNA:	1/4 Wave Mono-pole Omni Directional
RELAY CAPABILITY:	Automatic
ERROR DETECTION & CORRECTION:	RSED 31-15 (Reed-Solomon Encoder/Decoder)
TACAN PERFORMANCE:	Equal to or Better than the AN/ARN-84 TACAN

TABLE I (Cont'd)

AN/URQ-28 PERFORMANCE CHARACTERISTICS

OPERATIONAL MODES:	
MODE 1	Maximum A/J
MODE 2	Nominal A/J
MODE 3	Minimal A/J
MODE 4	No A/J
TDMA/TACAN:	Simultaneous TDMA and TACAN Operation or Manual Selection of Either TACAN or TDMA
TDMA NETS:	Up to 5 Simultaneous Nets in the Same Geographical Area on a Time Slot by Time Slot Basis

Passive Operation:

Terminal is capable of receiving and processing messages for up to 12 hours without time updates.

Block Diagram Description

A simplified block diagram of the AN/URQ-28 Receiver-Transmitter is shown in Figure 2.

Two independent super heterodyne receiver chains are incorporated in the terminal. These receivers can be used for two simultaneous TDMA channels or one simultaneous TDMA and one TACAN channel. The TDMA channels are single conversion and the TACAN channel employs dual conversion. The TDMA channels employ RF gain, tunable RF filters, wide-band double balanced mixers, linear-limiting IF amplifiers, and programmable surface acoustic wave (SAW) correlators. TOA and data detection employ both automatic and computer controlled (adaptive) threshold levels.

The Transmitter Exciter accepts the baseband digital signals from the Signal Processor and converts it into (CPSM)Controlled Phase Shift Modulation, which is a modulation technique similar to Minimum Shift Keying. This signal is then converted to the desired "L" Band output frequency by mixing the IF signal with an appropriate LO output from the Frequency Synthesizer. The output of the Exciter is then amplified to 800 watts peak power in the wideband solid-state RF Power Amplifier.

The Frequency Synthesizer is a dual phase-locked digital synthesizer capable of switching from any one "L" band frequency to any other in microseconds. The unit also provides outputs for the 126 TACAN channels.

The Signal Processor provides TDMA synchronization, ranging position location, data processing, error detection and correction as well as all terminal housekeeping functions. The signal processor incorporates an all LSI digital computer with 28K 16 bit words of solid state memory.

The AN/URQ-28 measures range in either the TDMA or TACAN modes or both. TACAN ranging is accomplished to ground or surface beacons (or to other aircraft in the air-to-air mode) with an accuracy equal to the AN/ARN-84 TACAN.

TDMA ranging is based on the establishment of highly accurate relative time synchronization among all community members. TOA (Time of Arrival) of received transmissions then becomes an accurate measure of slant range between the transmitting and receiving members. All community members are capable of making this range measurement on a transmission by one member.

Terminal Software

The software for the AN/URQ-28 Tactical Terminal is divided into three major functions, I-TDMA Processing, II-TACAN Processing, and III-Relative Navigation Filter Processing. The TDMA processing consists of synch processing, TDMA housekeeping, message/slot processing and message/port processing. The TACAN processing consists of an envelope filter, phase reference tracker, DME acquisition and track, interrogation logic, display processing and slot processing. The Relative Navigation processing consists of source selection and processing, navigation port interface processing, and the relative navigation Kalman filter.

I. TDMA Processing

The first TDMA Processing Function is the Synch Processing. This consists of the Round Trip Timing (RTT) synchronization filter and the various functions for controlling RTT reply, control of the time base, coarse sync and net time updates. The first step is coarse synchronization. This processing computes and controls the set up of the terminal to enter the JTIDS net. Based upon operator entered values for the Time of Day (TOD) and estimates of the accuracy of this TOD, the program computes a look ahead time. Based on this value of time, the hardware is preset for reception of a transmission in the future. When this transmission is received, coarse synchronization is achieved and fine synchronization is started. Fine synchronization is achieved either by RTT or by passive means. If RTT is used then the RTT source selection processing chooses the recipient of the RTT message and the RTT synchronization filter processes the RTT replies to estimate the clock bias and frequency drift. If passive synchronization is chosen, then clock bias

I. TDMA Processing (Cont'd)

and frequency errors are estimated by the navigation filter. The time base control software provides the necessary processing to control the time base and to apply the correction from either the synchronization filter or the navigation filter to the clock itself.

The TDMA housekeeping processing consists of the Net Control Processing, Test Message Control, TDMA Performance Monitoring, TDMA Receiver Threshold Control and Mode Control Processing. The Net Control Processing associates a TDMA net number and communication function with each slot time. Since the terminal is capable of switching between nets on a slot by slot basis, the net number, type of data, and whether it is a receive or transmit slot must be designated for each slot.

The next set of processing functions are called the Message/Slot Processing group. These functions are concerned with the transmission and receiption of messages on a slot by slot basis.

The first of these is the Received Message Screening and Routing. This routine provides the primary message interface between the software and hardware for all received messages. Not only does it control the data from the RF port, but it also outputs the data to the fixed format or Free Text Port.

The JTIDS program has generated an Interim JTIDS Message Specification (IJMS) which defines various types of messages. The fixed format messages are called Type I messages and the function of the Type I Message Processing provides the identification, classifying processing of these messages.

One particular class of Type I message is identified as the P-message. The P-message is the position and status message in the (IJMS) catalog and consists of the P1 (Airborne), P2 (Ground Station), and P3 (Ship). The P-message processing formulates the P-message for inclusion in the transmission queue at the proper time. This message is composed of both data supplied external to the terminal and data generated internally. It is also the function of this processing to select the desired transmission slot and to time extrapolate the navigation data to the beginning of the slot.

In addition to generating the P-message, the terminal also controls the relay of messages and generation of automatic machine acknowledge messages. The relay processing modifies the relay indicator bit in the message and supplies the transmission queue with the appropriate data to enable the transmission of the relayed data in the proper slot. The machine acknowledge processing routine provides the interface between the Type I message classification and the transmission queue. If a machine acknowledge type message is received, this processing determines if a reply is required and if the terminal mode permits sending one.

The Test Message Generating function has the responsibility of controlling the transmission of test messages and of generating the data for the test message.

RTT interrogation generation is a processing function which determines when an RTT message is required and permissable and designates the address of the unit to which such an interrogation is sent. This function is activated only if the terminal is in the active (RTT) synchronization mode.

The final message/slot processing function is the transmission queue processing. This function selects messages from the various sources within the terminal and from the external ports for placement in a queue.

The fourth block of software is designated the Message/Port processing. This processing is divided into four functions which are the free text port input, free text port output, fixed format port input, and fixed format port output.

II. TACAN Processing

This processing is divided into seven functions. These are the envelope filter, phase reference tracker, DME acquisition, DME tracker, interrogation logic, display processing, and time slot processing.

The envelope filter is used to estimate the phase of the TACAN 15Hz/135Hz composite signal. The estimation procedure is mechanized via a Kalman Filter. The estimated states are the in-phase and quadrature components of the 15Hz and 135Hz signal, the average DC level and additional estimated data with respect to the beacon.

The DME Acquisition Processing is utilized to provide an initial value of aircraft range for the DME Tracker. This process is a search procedure that separates the correct interrogation replies from all other squitter pulses emitted by the beacon. There are two modes of operation, the search mode and the monitor mode.

The DME Tracker is used to make fine estimates of range from the aircraft to the beacon and to maintain beacon track after coarse range acquisition has occurred. This is done with a two state Kalman filter.

The Interrogation Logic provides for the randomization of the times at which interrogations are transmitted. The Display Processing is used to compute the true bearing of

II. TACAN Processing - (Cont'd)

the aircraft and to supply digital range and bearing to the Display. The time slot processing incorporates all data buffering requirements for the collection of information needed for the track filter, envelope filter and phase reference tracker.

III. Relative Navigation Processing

This processing is divided into source selection and message processing, navigation port interface processing, and the relative navigation filter processing.

The relative navigation source selection and message processing supports the relative navigation through selection of suitable data for grid and geodetic processing. This function extracts the data required from the incoming P-message and provides the extracted data to the relative navigation filter processing.

The navigation port interface processing controls the interchange of data between the relative navigation filter and the dead reckoning equipment which is interfaced with the JTIDS Terminal.

The relative navigation filter processing function provides and maintains hybrid grid and geodetic navigation data.

Filtered grid and geodetic position and velocity, as well as best estimates of dead reckoning attitude errors and TDMA clock errors in passive sync mode are developed from the weighted mixing of dead reckoning inputs, TDMA time-of-arrival (TOA) data, and geodetic or grid position source reports derived from received P-messages. Four types of observation data processing are provided which include use of TOA data for improved grid navigation, TOA data for improved geodetic navigation, geodetic and grid data for offset geodetic updating, and geographic position update data provided from the nav port.

As an example, to process single range measurements, a range estimate is formed from the reported position and the estimate of own position by

$$\hat{R} = \left[(\hat{X}-X')^2 + (\hat{Y}-Y')^2 + (\hat{Z}-Z')^2\right]^{1/2} \qquad (1)$$

where x, y, z are the estimate of own position and x', y', z' are the reported position of the other member. A range residual is defined as the difference between the estimated and the measured range

$$\varepsilon_R = \hat{R} - R_M \qquad (2)$$

position error components are calculated by multiplying the range residual by estimates of the direction cosines

$$\varepsilon_X = \varepsilon_R \left[\frac{\hat{X}-X'}{R}\right] \qquad (3)$$

$$\varepsilon_Y = \varepsilon_R \left[\frac{\hat{Y}-Y'}{R}\right] \qquad (4)$$

$$\varepsilon_Z = \varepsilon_R \left[\frac{\hat{Z}-Z'}{R}\right] \qquad (5)$$

This processing of range measurements is performed by a Kalman Filter which has been mechanized to optimally mix the TOA observation data from the TDMA system with the dead reckoning navigation data.

AN/URQ-28 Relative Navigation Capability

The Relative Navigation capability is provided by making use of onboard sensors. In operation, the dead reckoning equipment senses vehicle motion and the computer calculates position by integration. A position report from a remote unit is received via the JTIDS link and the TOA of this message indicates the range to the remote unit. Using this information, a correction to the estimate of own position is calculated and applied. In turn, a position report is transmitted via JTIDS for use by other members.

The coordinates frame used for position reporting in the Relative Grid is a tangent plane grid, a rectangular grid tangent to the earth at the origin, which is nominally stationary and north oriented.

The origin of the grid is located arbitrarily, but all members accept the location identified by one member of the community designated as controller. From the transmitted data, the apparent relative position of each community member is determined. If each member's self-contained grid is co-aligned, this apparent relative position will be the same as the relative positions measured by the radio ranging.

The difference between the measured and apparent relative position provides the error signals which enable each member to align his self-contained grid to community grid. Thus, members each have accurate navigation in the tactical grid in spite of geographic navigation errors.

AN/URQ-28 RELATIVE NAVIGATION CAPABILITY - (Cont'd)

Geographic position is not affected by this processing of range updates unless one of the members of the community has an accurate geographic fix or some source of geographic data such as LORAN, OMEGA, or satellite receiver. If such geographic data is available, the tactical grid can be used to distribute the data to other community members.

Each community member reports his position at the time of the range measurements, thus providing the location of the datum point. If no range measurements are available, the system navigation is based on the dead reckoning outputs.

The conventional approach to the use of range measurements is to correct position error by taking simultaneous measurements to separate datum points and calculating a complete fix (i.e., position measurements in three dimensions). The Relative Navigation that Singer-Kearfott has implemented processes range measurements individually, making best use of the available information in a situation where the geometry is continually changing. When a range measurement is available, the position estimate is corrected in one direction, giving a partial fix. When multiple measurements are available, they are processed independently, yielding corrections in whatever direction geometry permits.

AN/URQ-28 MECHANICAL DESIGN

The AN/URQ-28 JTIDS Tactical Terminal has been designed in a 1 1/4 ATR configuration in order to accommodate a majority of military applications with minimum volume and weight impact. Figures 3, 4, 5, and 6 illustrate the design.

Physical characteristics of the terminal are as follows:

Length	19.526 inches
Height	7.625 inches
Width	12.656 inches
Weight	85 lbs
Heat Dissipation	800 watts

This high density packaging concept for military avionics has been developed and produced by Singer-Kearfott for B1, (Bomber), JA-37 (Swedish Fighter), as well as the JTIDS.

The design employs a stacked and clamped cards mounting technique in which the card frames form an integral "box-like" structure for the overall assembly, thereby eliminating the need for an external chassis. Four bars are used to clamp cards between a front connection panel and a rear support panel. Removal of Signal Processor Logic Modules and Power Supply Modules is effected by loosening four rear nuts and lifting the bars. Bars on the opposing side form a sub-chassis by providing a structure for retaining separate Signal Processor and Power Supply female connector plate/motherboard assemblies and for keying and guiding modules. Future growth can be accommodated by extension of the motherboards with flexprint wiring interconnections and appropriate bar length extension. The other half of the case is occupied by the RF sub-assembly.

The RF modules are also held in place and clamped with bars, however, wiring interconnections are made with bus bars, coax, and cables. No motherboards are used on the RF Sub-Assembly. Disconnection of connectors and module removal is done in a vertical direction as contrasted to the sidewise removal on the opposite side.

The front panel is a solid aluminum casting and all electrical connections are made via the front panel. The rear panel is made in two pieces which are pinned together to permit disassembly for maintenance and access to the motherboards. Figure 7 defines the overall dimension of the AN/URQ-28 Tactical Terminal. Signal Processor and Power Supply module connectors incorporate the standard NAFI blade and tuning fork design utilizing pins captured in individual nylon inserts.

The terminal employs hollow modules as the heat exchanger mechanism. Cooling air passes directly behind components through a parallel plane gap formed by circuit card lamina cemented to each side of a module frame. The module frame contains an inlet and exit port. By stacking and clamping the modules, their inlet and exit ports form an air distribution plenum.
Air entering at the rear panel flows through to the inlet plenum and is distributed to modules in a parallel fashion, each module therefore, being exposed to the same inlet air temperature. This is in contrast to more conventional series flow design where each module is exposed to successively higher heat exchanger interface temperatures. After cooling air leaves the module, it enters the exit plenum which directs the air to exit ports located at both rear and front panels.

Extensive thermal analysis and testing has been performed to establish the terminal thermal characterisitcs. Module inlet ports are adjusted to "bleed" the desired amount of cooling air based on the dissipation within the module.

The terminal has been designed to meet MIL-E-5400 Class 2X environment requirements for fixed wing aircraft and Class I for helicopter applications.

The terminal is designed for "hard mounting" or for use with vibration isolators and a mounting rack. Cooling air is provided either from an aircraft conditioned air supply or by blowers included in the mounting rack.

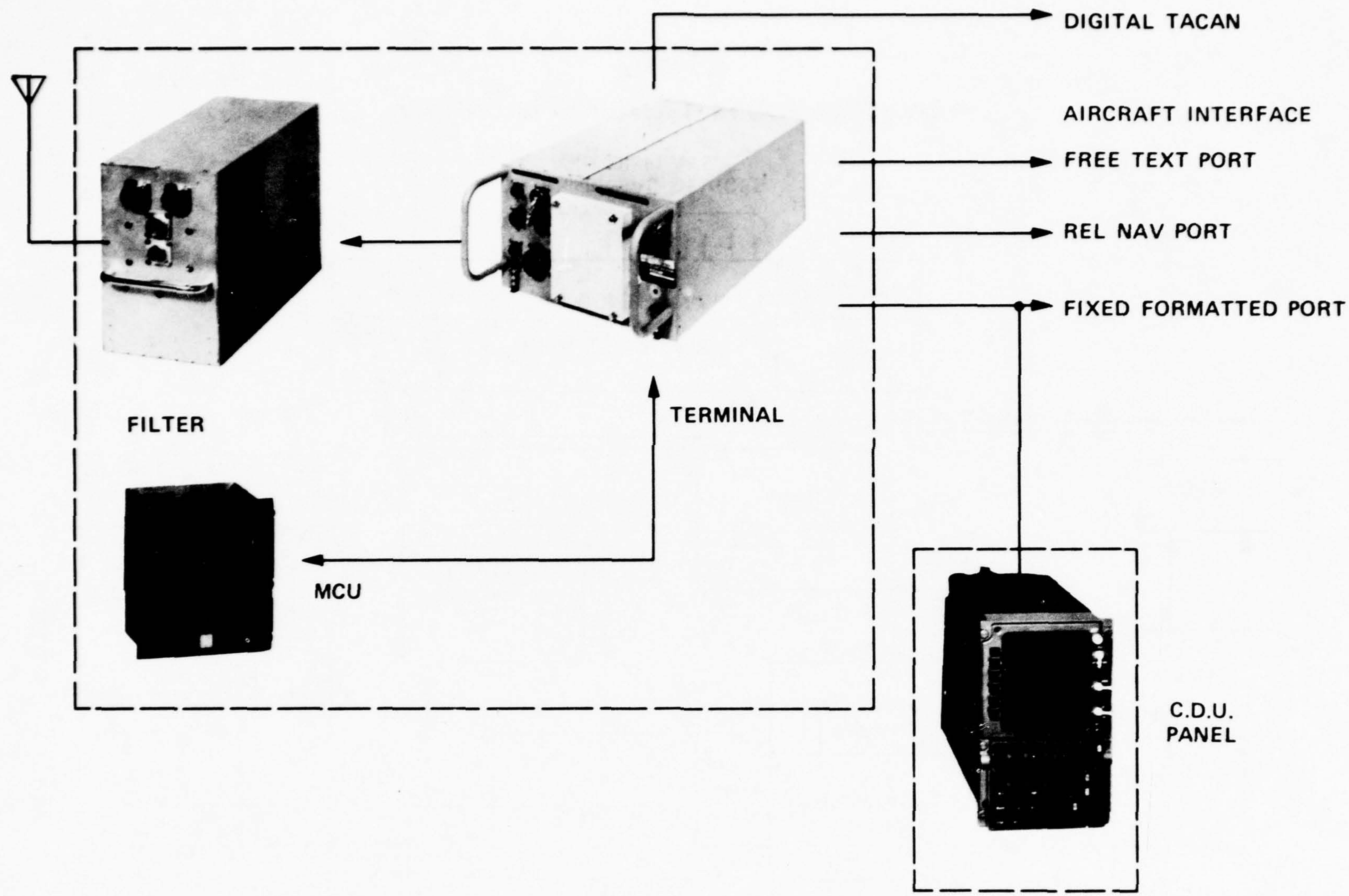
AN/URQ-28 MECHANICAL DESIGN - (Cont'd)

Each logic module can accommodate up to 120 14 pin flat-packs, 60 on each side or an equivalent mix of MSI/LSI packages with as many as 48 leads.

The Power Supply and RF module frames are also 5" high and 7.415" long, however, their widths vary from 1.275" to 2.450". The basic construction is similar to the logic modules, however, the cooling air passage incorporates finned material to increase surface area. These modules incorporate circuit cards and RF, IF, and video components. The RF power amplifier is 5.750" wide and is made up of 4 submodules bolted together. RF shielding is accomplished with module covers and compartmentizing. All RF circuits are completely shielded from the air passages. Adequate removal of heat in a Class 2X environment is an outstanding feature of the AN/URQ-28.

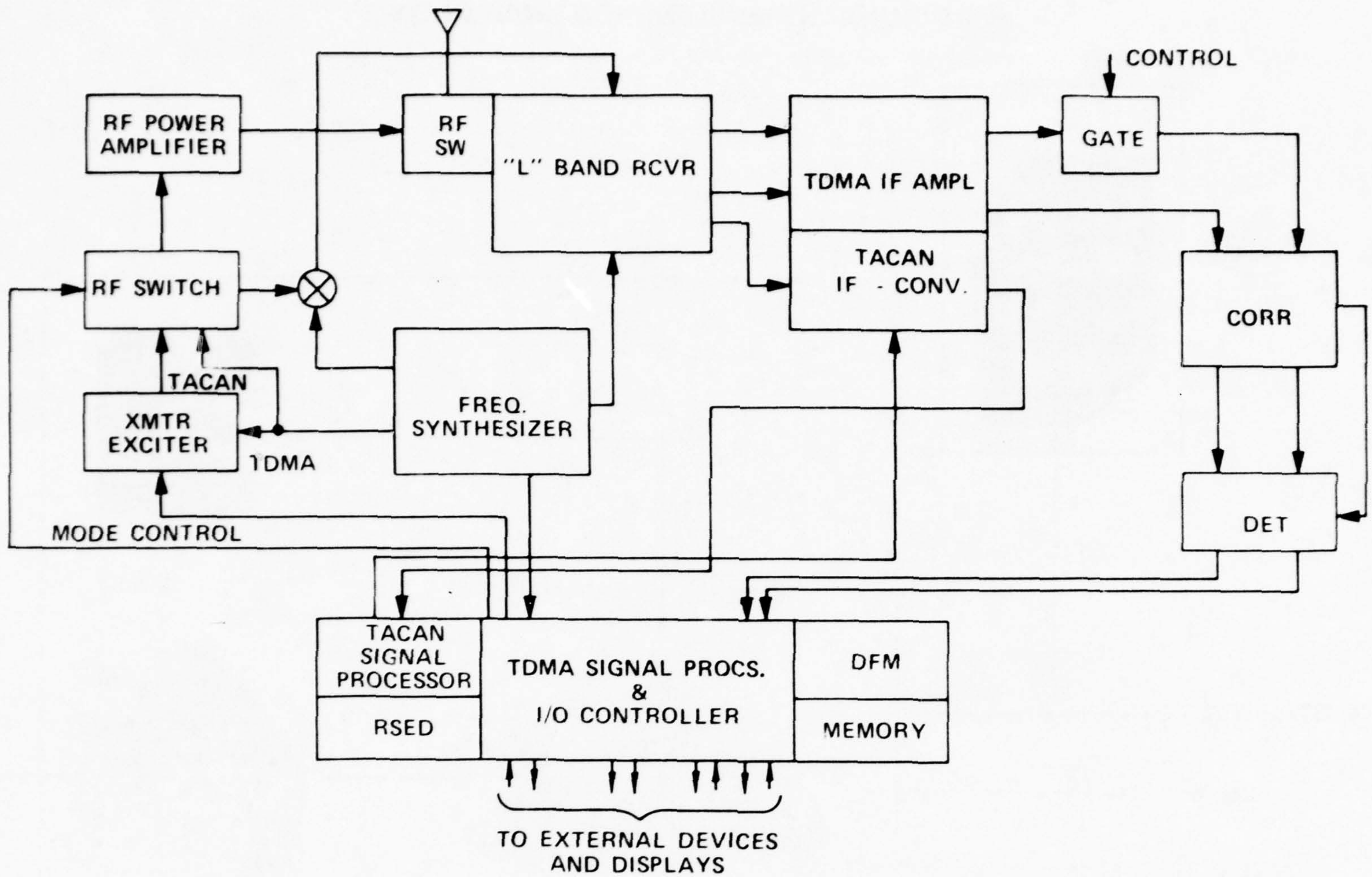
The AN/URQ-28 Mode Control Unit (MCU) is shown in Figures 8 & 9. This is a cockpit mounted module containing functional controls for the terminal. TDMA mode control and TACAN channel select and control occupy the upper half of the panel, while the lower half contains the necessary controls for the terminal Secure Data Unit (SDU).

The AN/URQ-28 JTIDS Tactical Terminal is one of the first truly integrated aircraft communication and navigation systems and, as its capabilities are realized and understood by military users, the terminal will permit significant improvements in tactical operations. The multi-mode capability of the JTIDS will permit a broad spectrum of mobile and fixed users to participate in the community.



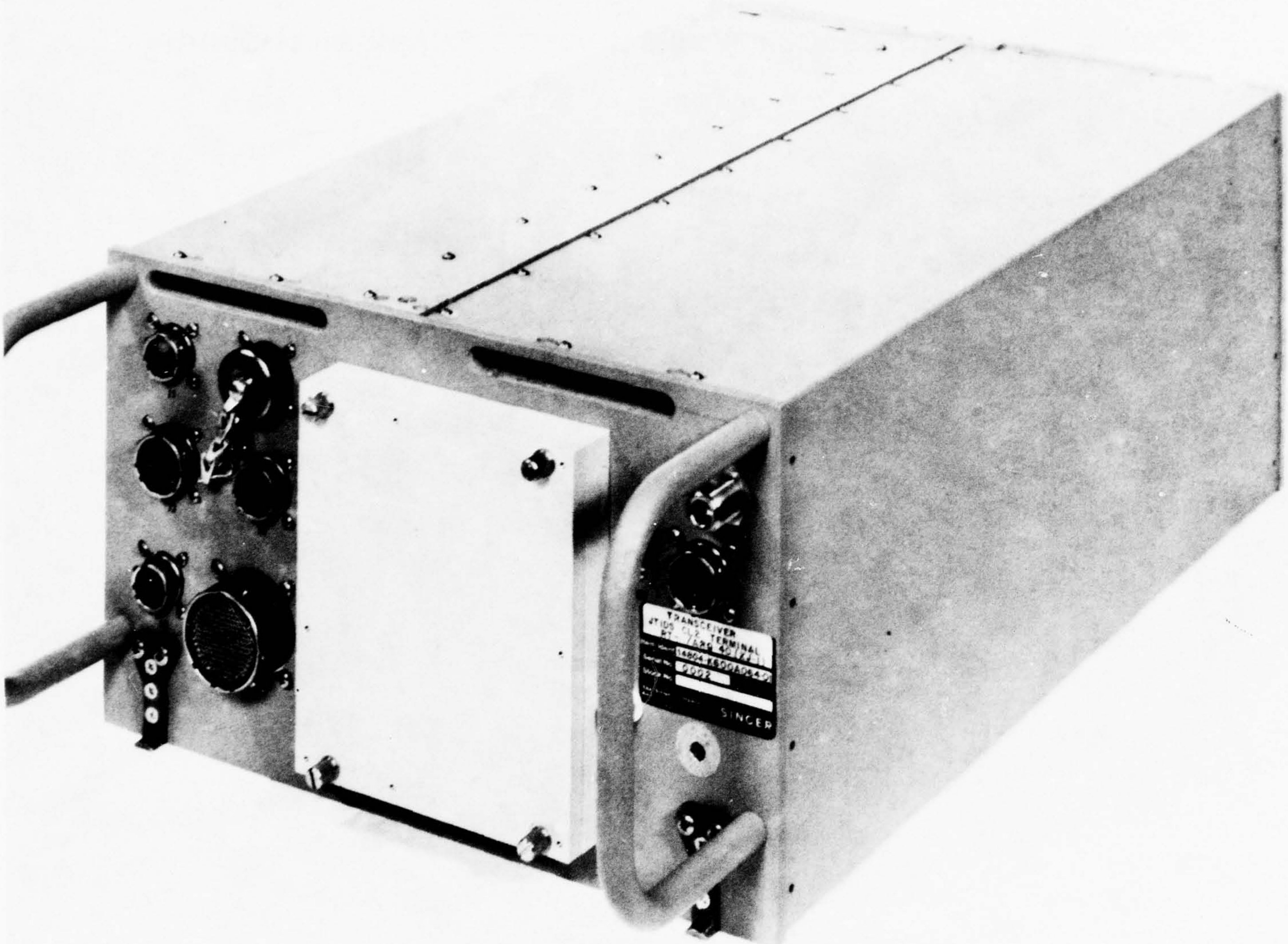
EQUIPMENT CONFIGURATION (AN/URQ-28)

Figure 1



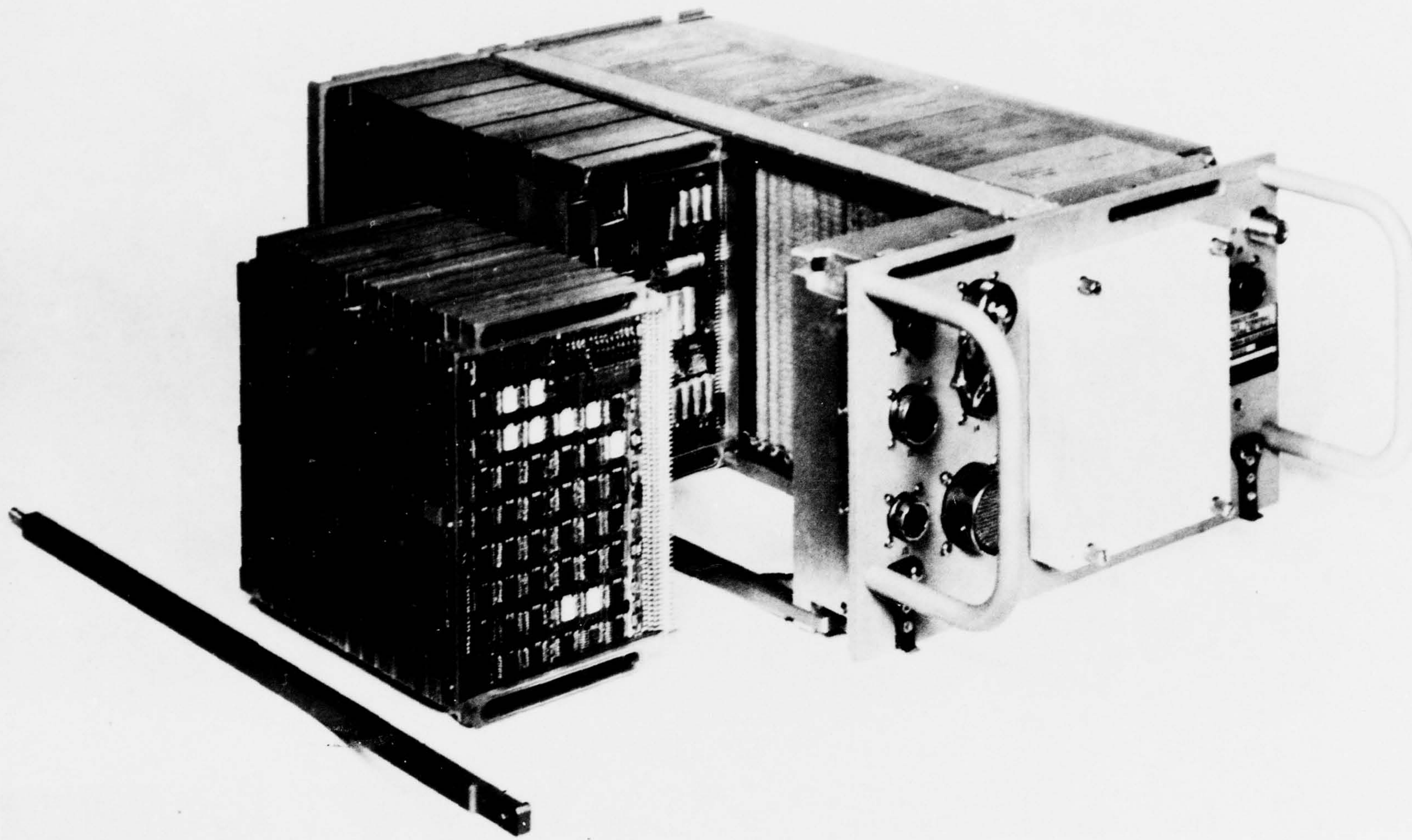
AN/URQ-28 SIMPLIFIED BLOCK DIAGRAM

Figure 2



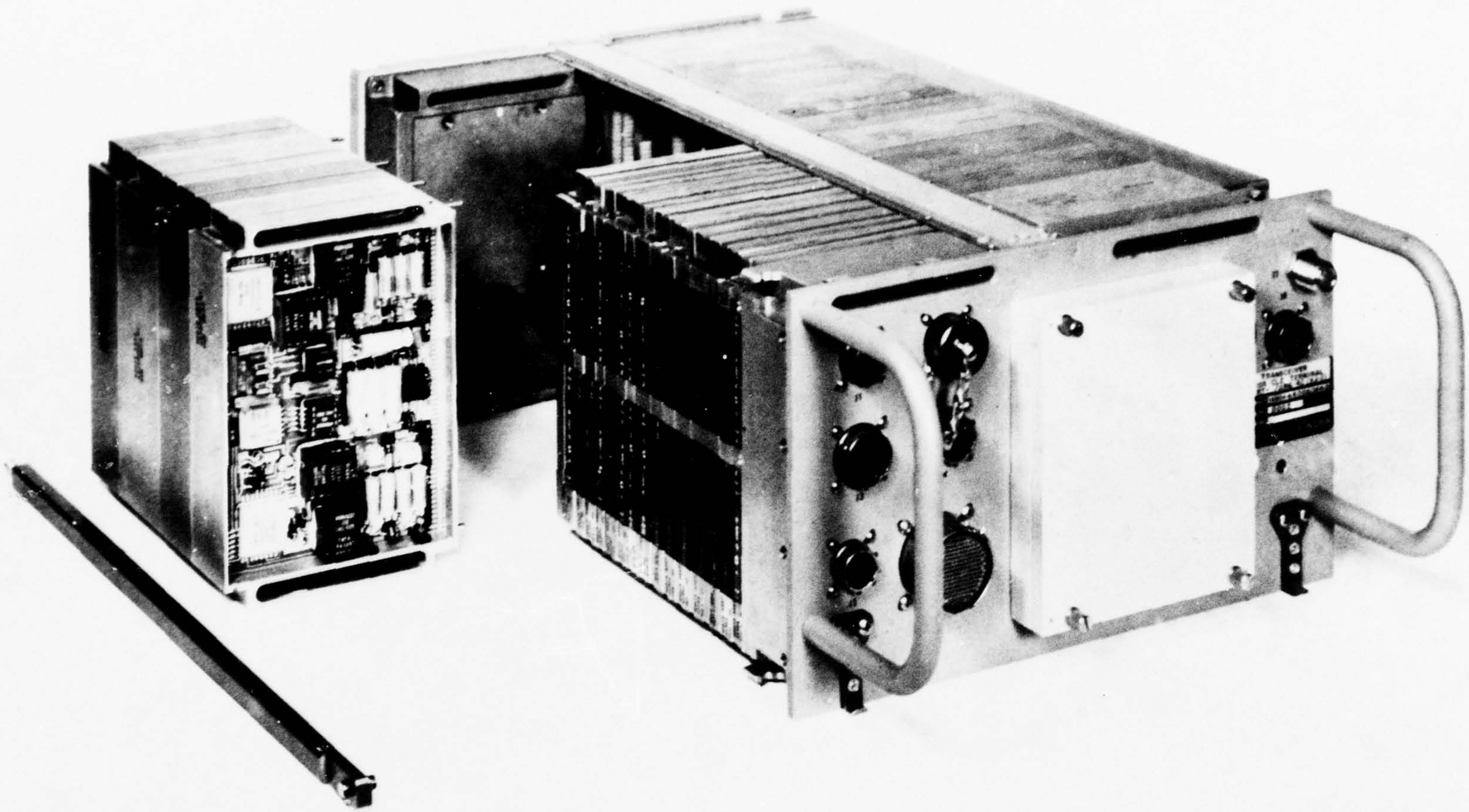
AN/URQ-28 MECHANICAL DESIGN RECEIVER-TRANSMITTER-PROCESSOR

Figure 3



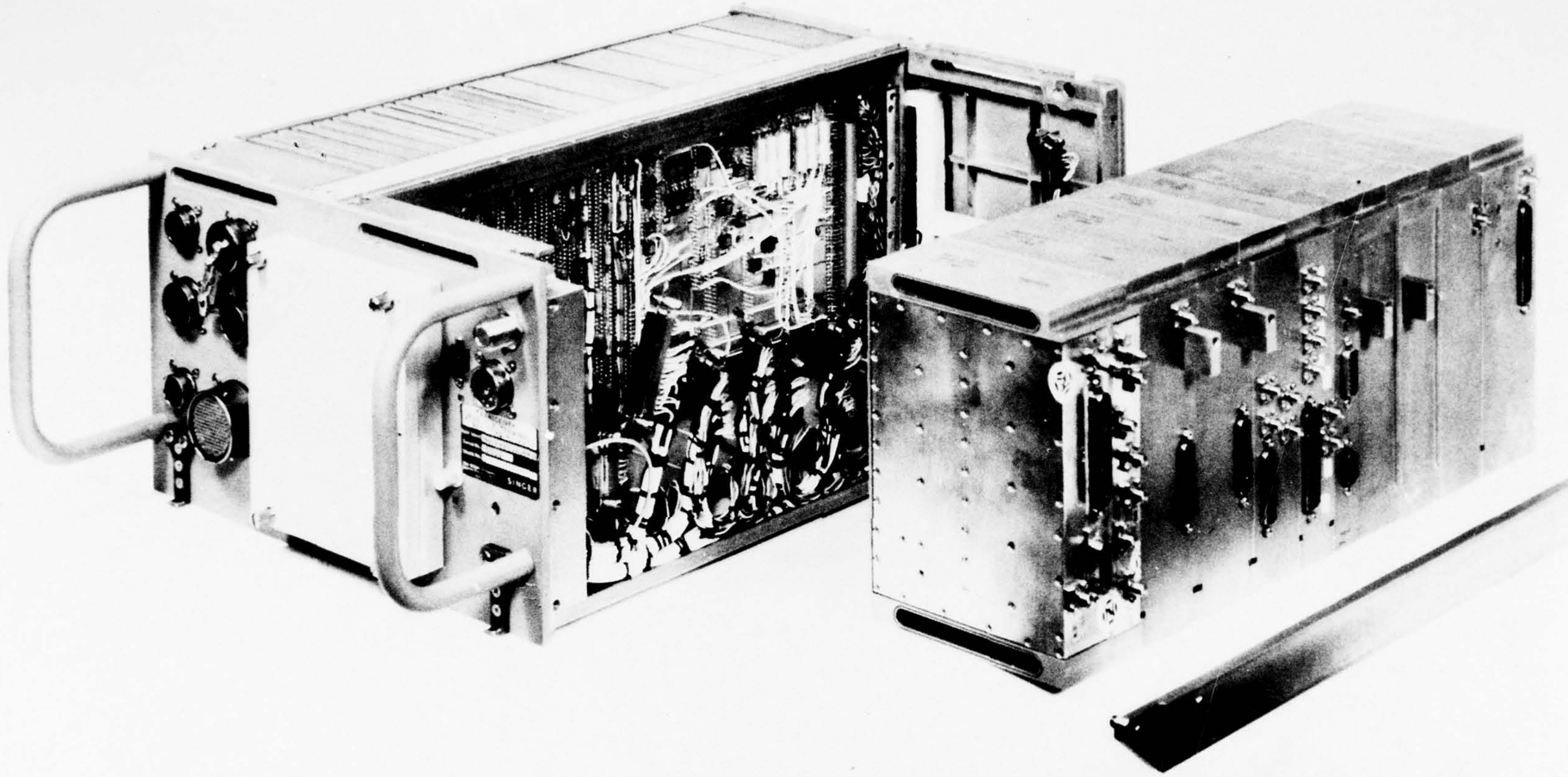
AN/URQ-28 MECHANICAL DESIGN SIGNAL PROCESSOR SUBASSEMBLY
(15 CARDS EXTENDED)

Figure 4



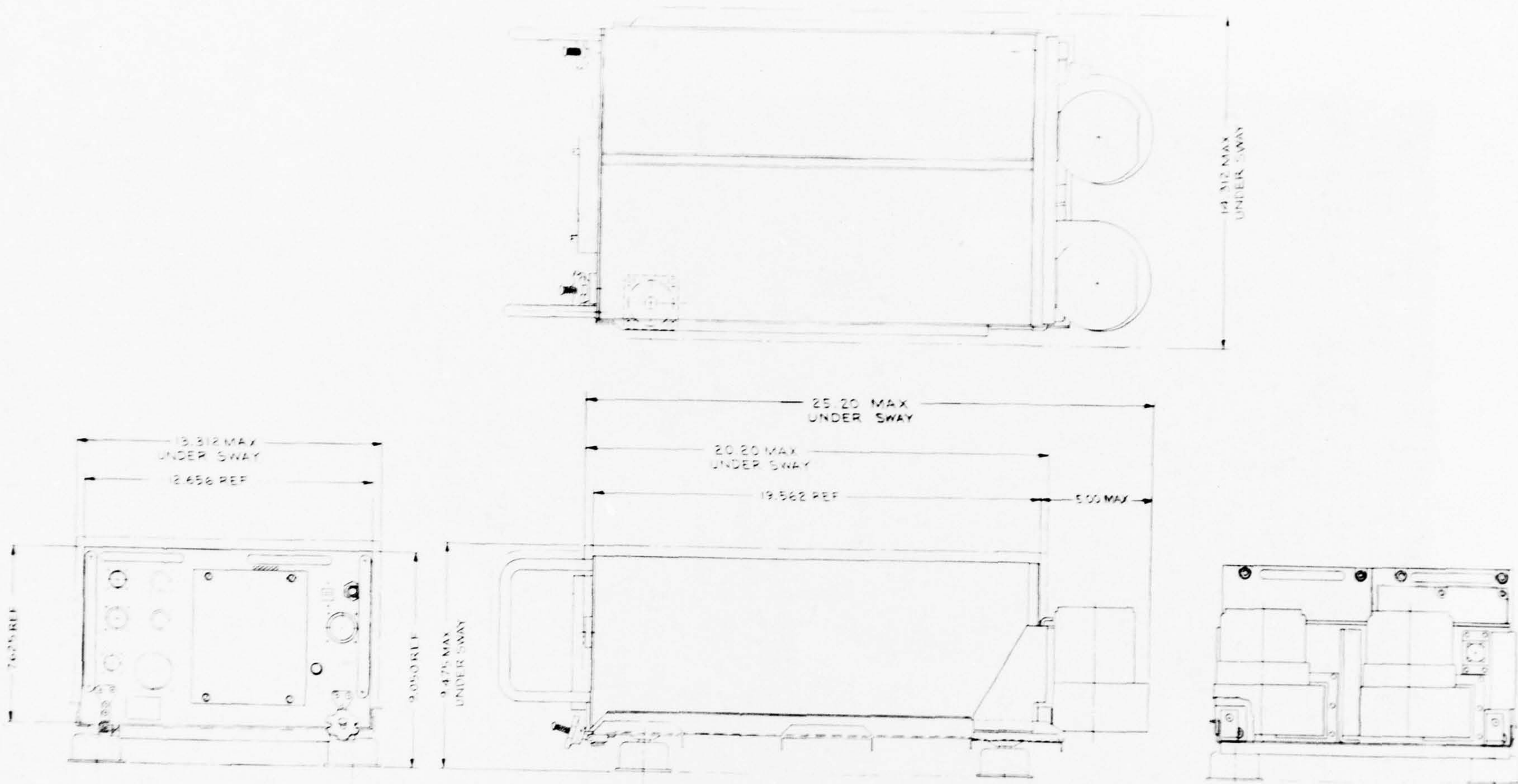
AN/URQ-28 MECHANICAL DESIGN POWER SUPPLY SUBASSEMBLY
(4 MODULES EXTENDED)

Figure 5



AN/URQ-28 MECHANICAL DESIGN RF. SUBASSEMBLY
(7 MODULES EXTENDED)

Figure 6

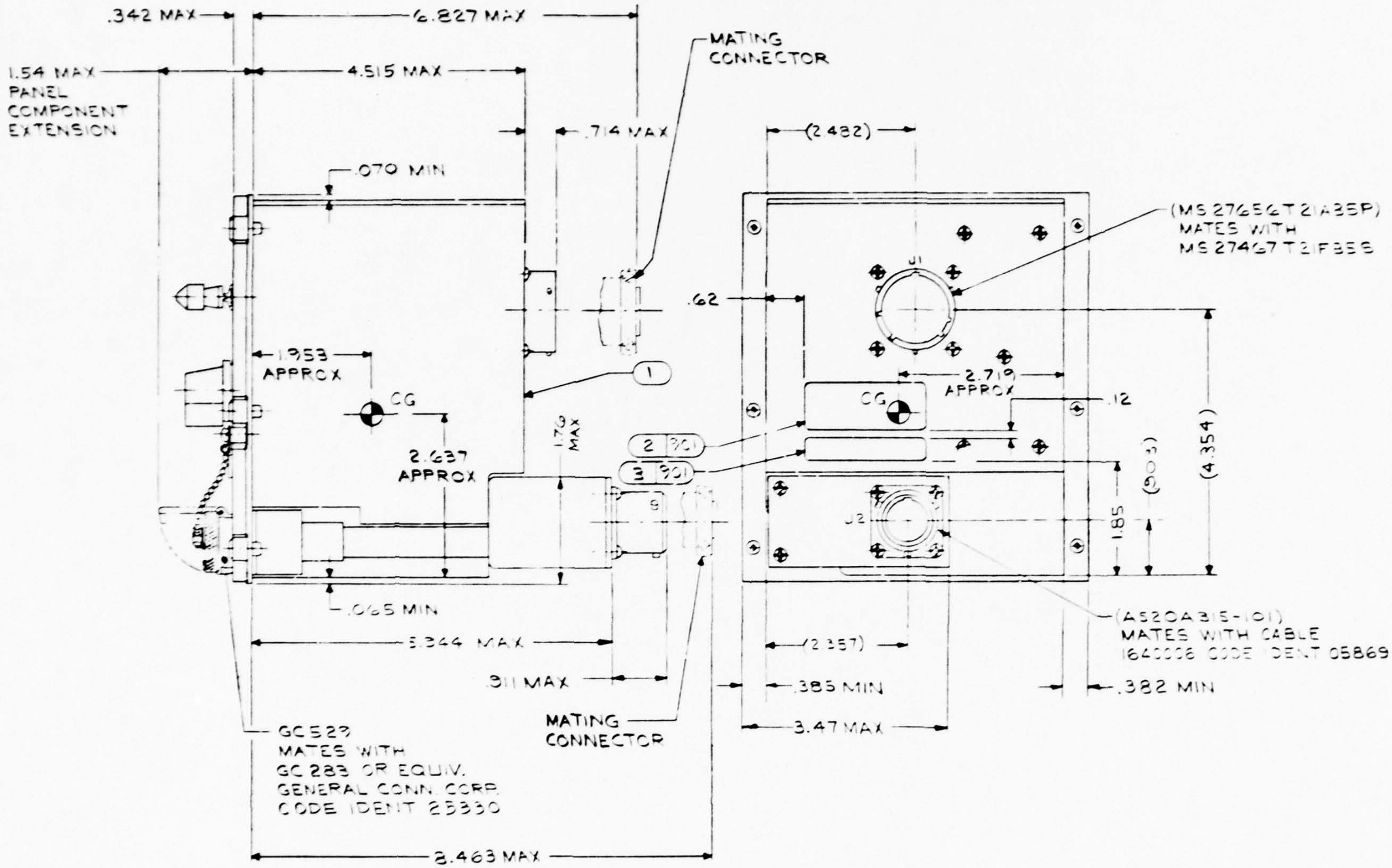


AN/URQ-28 TACTICAL AIRCRAFT TERMINAL OVERALL DIMENSIONS

Figure 7

JTIDS MCU

Figure 8



MODE CONTROL UNIT (MCU)

Figure 9

JTIDS EXPENDABLE/LOW COST TERMINAL DEVELOPMENT

by
P.D. Danker
R.J. Giannini
I.S. Glickstein
IBM Federal Systems Division
Owego, New York 13827 USA

INTRODUCTION

1.1 OVERVIEW. As depicted in Figure 1-1, many potential DoD applications exist for very low cost, expendable, JTIDS-compatible terminals, which include: (1) weapon and RPV guidance and control, (2) remote sensors and beacons, and (3) range instrumentation. In these and other situations, it is advantageous to utilize the basic JTIDS waveform for near-transparent communications compatibility with other JTIDS Class 1, 2, or 3 terminals. Expendable Terminals (ET) can be viewed as extensions for the other JTIDS Class terminals where (1) expendability and unmanned operations are required and (2) system communications/control resides with the JTIDS Class 1, 2, or 3 terminal.

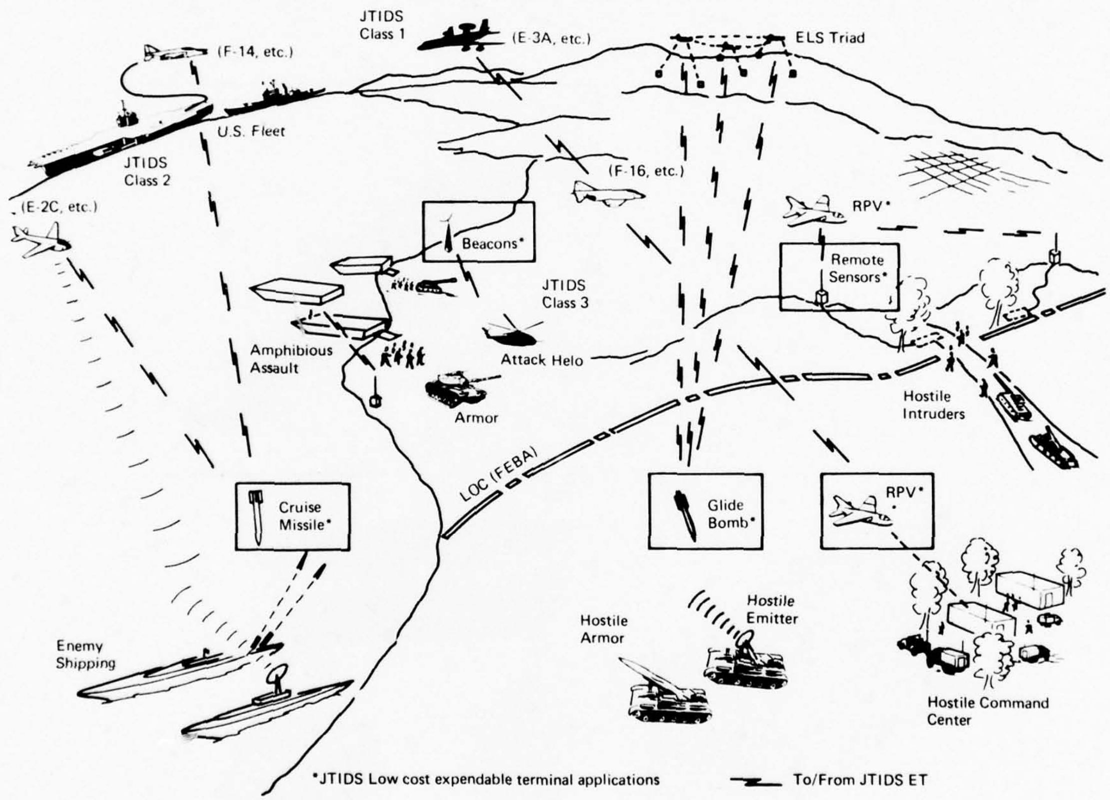


Figure 1-1. JTIDS Scenario

The major ET design emphasis is cost/performance. The system designer must search for and recognize opportunities to eliminate or simplify certain JTIDS features not required for a specific use. IBM has gained an insight into complexity vs JTIDS features tradeoff techniques as a result of participating on a number of recent programs, which include: (1) production of expendable terminals for the HOBO glide bomb, (2) jam resistant drone terminals, (3) in-house data link system development and testing of a brassboard JTIDS ET, and (4) a JTIDS Class 3 terminal conceptual study contract. This paper discusses cost/performance issues, JTIDS features vs complexity tradeoffs as they affect ET functions, impact of technology infusion, and production base considerations, using the IBM JTIDS ET brassboard design as a development benchmark.

The major functions of a JTIDS terminal are: (1) receive and correlate the JTIDS waveform, (2) decode and process the messages, (3) interface with the vehicle and/or sensor in the user's installation, and (4) format and transmit unique JTIDS messages. ET cost/performance vs complexity tradeoffs affect one or more of these functional areas. For instance, in a number of potential applications, there is no need for an uplink or a downlink capability, or a need for only a few control or status bits, rather than a

complex message. In these cases, the receiver or transmitter, and associated processing, may be eliminated or simplified.

In some applications, less complex implementations are possible in the areas of anti-jam (AJ), security, data rate, and net entry. These and other techniques discussed in subsequent sections of this paper can impact production terminal cost without measureably affecting mission performance. In fact, overall performance may be cost effective as a result of the size, weight, and power reduction, and the reliability increase due to a less complex terminal.

A large production base is also necessary to low unit cost. If common/base functional modules can be produced, one can benefit from technology infusion in the form of custom large scale integration (LSI), special microprocessors, dense RF packaging, matched filters, etc. Each of these areas of production cost savings requires research and development expenditure and non-recurring start-up expenses that should be justified only when spread over a large production run. Cost goals in the range of $5 to $10K may be met for production quantities of 10,000 or more units. LSI and dense packaging are also key to meeting the stringent weight goal of less than 10 pounds, the form factor goal of less than 500 cu. in., and the primary power goal of less than 50 watts.

1.2 IBM DEVELOPMENT BACKGROUND. IBM has been involved in the development of time of arrival (TOA) and distance measuring equipment (DME) systems for more than a decade. Over 200 expendable DME Weapon Guidance Subsystems (WGSS) for the HOBO/GBU-15 glide bomb (Figure 1-2) have been produced by IBM. Although the WGSS is not JTIDS-compatible, (since it was designed prior to JTIDS standardization) it does serve as a real-world example of a practical expendable terminal.

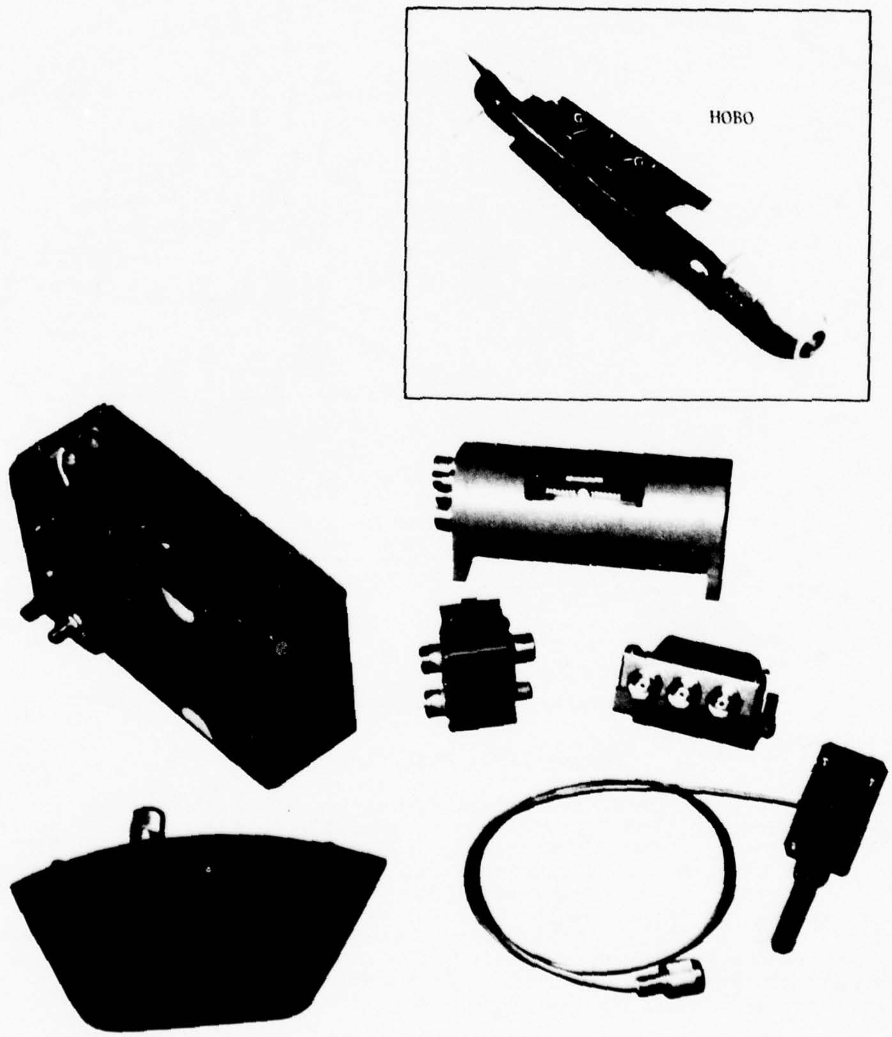
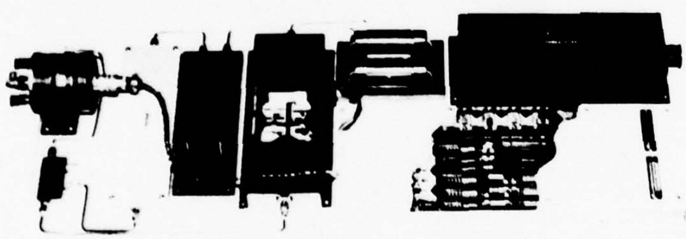


Figure 1-2. Expendable Weapon Guidance Terminal (WGSS)

The GBU-15/WGSS work lead to the Drone Formation Control System (DFCS) DME terminal development, for the White Sands Missile Range. This terminal development, in addition to providing precision DME location and control of drone vehicles, also integrated spread spectrum PN coding techniques into its L band waveform to aid in overcoming multipath and noise problems thereby enhancing overall DME system performance. As a result of the DFCS terminal development, unmanned beacon stations, range instrumentation systems, and the ABS-I terminal were delivered (shown installed in drones, Figure 1-3a). Further development effort for follow on range instrumentation applications resulted in the ABS-II terminal concept shown in Figure 1-3b.

IBM's JTIDS activity ranges from the AWACS terminal development (for which an IBM processor is supplied) to the ASIT terminal (for which IBM is the systems developer). In addition, a number of study contracts and in-house development programs have been performed in JTIDS and related technology areas. Of particular value to expendable terminal development are (1) the JTIDS Class 3 terminal conceptual study contract, (2) the JTIDS/JAWS development program, and (3) an in-house development of a JTIDS brassboard ET. The material presented in this technical paper is based on the experience gained from the above development programs.

Figure 1-3a. ABS-I DFCS Terminal Installed in WSMR Drones



Physical Characteristics	
Weight	35 pounds
Size	10 x 24 x 3 inches
Power	100 watts

Figure 1-3b. ABS-II Development Terminal

2. SYSTEM REQUIREMENTS AND DESIGN TRADEOFF

2.1 TYPICAL APPLICATION REQUIREMENTS. Five potential DoD application areas for JTIDS ETs are identified and will be used in this paper as examples for which typical design requirements may be derived. The following paragraphs describe each of these application areas.

a. Guided weapons and missiles - This class of ET applications covers a spectrum from DME-guided glide bomb, of relatively short range, to the tactical anti-ship cruise missile, with an over-the-horizon operational range. In general, the uplink messages to the missile or weapon are full-length, containing flight commands or position data for use by the missile in computing its own flight commands. Downlink messages from the missile or weapon will normally be very simple, or not required at all. Typical downlink messages may consist of JTIDS RTT (round trip timing) responses, and/or bits indicating weapon status/location.

b. Remotely piloted vehicles (RPVs) - range from simple/low cost Aquila type, used by the Army as artillery spotters and for photo recon tasks, to more complex RPVs. The simpler RPVs have functional requirements very similar to the

guided missiles and weapons, given above. The more complex RPVs may require a more complete downlink capability, for reporting reconnaissance data. The downlink messages for these applications will be full length JTIDS messages.

c. Beacons - are devices that continually transmit a coded signal, or that transmit a coded signal in response to a coded interrogation. They may be surface implanted or air-dropped, or they may be water located buoys. Beacons can serve as position references in such operations as amphibious assault, where beaches and approach lanes must be marked. Uplink messages to the ET may not be required. The beacon, in this case, would be preprogrammed to transmit at appropriate times. If response on interrogation were required, the uplink would most likely consist of a fixed-format message. This uplink message might be a few bits of information, to roughly update beacon knowledge of JTIDS time and to cause the beacon to respond. However, a longer coded message might be used to prevent the beacon from responding to unauthorized (e.g. enemy) transmissions.

d. Remote sensors - Remote sensor ETs could involve surface located intrusion sensors that are planted along a border or military line of contact. The sensors would detect passing enemy traffic, using pressure, noise, odor, infrared, or other sensor techniques. The number and type of detections may be recorded for later burst transmission by the ET, or may be immediately transmitted by the ET, either case using full message uplink capability. In the first case, the ET might be programmed to transmit at pre-selected times, or it might transmit only following a coded interrogation. In operational use, an RPV or manned aircraft would pass by the remote sensor at intervals and read out the stored data.

e. Range Instrumentation - Test ranges need to provide instrumentation to track aircraft, missiles, and other weapons during weapon system test and evaluation. ET type terminals will be installed on the test vehicles (a/c and/or weapon), and additional ETs will be located at appropriate surveyed location points. During range use, these terminals will exchange transmissions to determine the ranges between the test vehicles and the surface ETs for calculations of test vehicle positions. Depending upon the ranging filter mechanization (and the potential downlink of flight parameters from the test vehicle to the surface ETs) the up and downlink message requirements will vary. For example, the test vehicle terminal may be downlink only, and the surface terminals may utilize time difference of arrival (TDOA) to locate the test vehicle. Conversely, the vehicle terminal may be uplink only, the surface ETs would continually transmit, and a processor aboard the test vehicle would compute vehicle position.

2.2 DESIGN REQUIREMENTS. Table 2-1 summarizes the design parameters which must be considered to satisfy each of the above described potential applications. These design parameters are:

a. Uplink messages (to ET) - range from full message length to no uplink requirement. Thus it might be possible for the receiver and associated processing portion of the ET to be modular, and left unpopulated if not required. The requirement for security is probably minimal so no decryption processing is required. AJ (anti-jam) and ECC (error correction) ranges from minimum to moderate.

b. Downlink messages (from ET) - range from full message length to no downlink requirement. Therefore, the transmitter and associated processing should be modular. AJ, ECC, and security range from minimum to moderate. Thus, frequency agility, error correction encoding and encryption could be modular and associated processing added as these features are required.

c. Data rate - None of the uplink or downlink requirements are projected to exceed 43 time slots, or transmissions per second for the applications examined. Thus, the ET may be designed to utilize every third JTIDS time slot (a so-called JTIDS "set"). This will allow the error decoding and message processors to take three time slots to perform their function (nearly 24 ms) rather than the normal 7.8125 ms. This would allow a slower, lower power, digital processor to do the job.

d. Transmission range - The maximum surface to surface ranges are projected in the order of 5-10 km. This can be achieved with peak transmission power of about 50 watts, depending upon reasonable foliage and terrain conditions. 10 to 50 watts should be adequate for surface to air situations, (up to 100 km) and for the air-to-air cases (up to 200 km). These requirements depend upon jamming environment by hostile forces, and the degree of AJ/ECC features incorporated in the ET.

Table 2-1. DESIGN PARAMETERS FOR ET APPLICATIONS

APPLICATION

DESIGN ELEMENT REQUIREMENTS	GUIDED MISSILES, WEAPONS	RPVs	BEACONS	REMOTE SENSORS	RANGE INSTRUMENTATION
Uplink messages (to ET) -Max data rate** -AJ, ECC, security	Full msg. 20 Moderate AJ, ECC, min. security	Full msg. 20 Moderate AJ, ECC, min. security	Minor (or none) 0-43*** Minimum	Minor (or none) 0-43*** Minimum	Full, minor or none* 0-43*** Minimum, multi-path rejection
Downlink message (from ET) - Max data rate** - AJ, ECC, security	Minor 0-7 moderate	Full (or minor) 0-7 moderate	Full <0.01ǂ - 43*** min.-moderate	Full <0.01ǂ - 43*** min.-moderate	Full, minor, or none* 0-43*** minimum
Max range -surf/surf -surf/air -air/air	 - 50-100 km 50-200 km	 - 50-100 km 50-200 km	 5-10 km 20-50 km -	 5-10 km 20-50 km -	 - 5-50 km -
User I/O - serial or parallel digital	short distance (<10 ft)	short distance (<10 ft)	none	long distance (>100 ft)	long distance (>100 ft) for surface units
Mechanical - size (cu in) - wt (pounds) - environment - prime power	 <500 <10 Missile 24-28 VDC	 <500 <10 RPV 24-28 VDC	 Not critical Not critical Surface 24-28 VDC	 Not critical Not critical Surface 24-28 VDC	 <500 <10 Airborne & surface types 24-28 VDC (or 120/208 V 400 Hz)

* Application dependent ** Time Slot (or portion of TS)/s *** One JTIDS set
ǂ I.e., 1 transmission per minute

e. User I/O - The ET must interface with the vehicle/autopilot in the missile/RPV/airborne range instrumentation systems, and with on-board sensors in the RPV and missiles. This subsystem function could utilize standard serial or parallel digital interfaces. Remote sensor/ETs must interface with the sensors themselves, which may be more than 100 feet distant from the ET. This may require a modulated FSK or other special interface so that ordinary wire (twisted pair) can be used. Similarly, the surface ETs in the range instrumentation case may utilize wire links to data reduction computers.

f. Mechanical requirements - Size is critical for the airborne applications. A value of the order of 300-500 cu. in. (e.g. 4" x 6" x 12' or 6" dia. by 10") is about as large as can be accepted in most missile, weapon and RPV applications.

Weight is also critical for the airborne applications. 8-12 pounds could represent some 10-20% of the total weight of some low cost RPVs.

Environmental requirements vary from airborne to surface. Since the ET is "expendable" and may be required to operate for relatively short periods of time in some of the applications, it may be possible to stress the unit by operating above normal temperature limits, etc.

Prime power is 24-28 vdc in all cases, except some range instrumentation situations where 120/208v, 400 hz may be preferred.

g. Cost - Insure affordable cost range ($5k-10k) by continuously trading performance/function vs cost. Utilize design to cost techniques to enable achieving cost goals.

In summary, the requirements vary considerably from application to application. One common thread is the use of the ETs in unmanned applications, eliminating requirements for displays and controls, and possibly reducing reliability and backup provisions. A second possible common thread is the elimination of high AJ/ECC requirements, and the high data rates or high transmission power levels.

2.3 JTIDS FEATURES. The JTIDS waveform and message discipline afford a great deal of flexibility in terms of operational modes, anti-jam features and data handling for a variety of applications. Three classes of JTIDS terminals have been defined:

(1) Class 1 terminal - developed for command and control applications.

(2) Class 2 terminal - for use on tactical aircraft/shipborne platforms, and

(3) Class 3 terminal - low cost, light weight terminal applications including a manpack.

These classes of terminals are characterized by varying degrees of performance and cost. The JTIDS ET will be less complex and cost less than the Class 3 terminal, and provide the specialized performance needed to satisfy the requirements for the unmanned/expendable type of applications discussed previously.

Table 2-2 summarizes the JTIDS features needed to support the various ET applications. An entry in the table indicates that a feature is required (or needed more than a minimal amount). A discussion of these features follows:

Table 2-2. JTIDS FEATURES FOR VARIOUS ETs

JTIDS FEATURES	GUIDED MISSILES, WEAPONS	RPVs	BEACONS	REMOTE SENSORS	RANGE INSTRUMENTATION
TDMA	X	X	X	X	X
Interoperability	-	-	X	-	X
Net Synchronization	X	X	-	X	X
AJ Modes	X	X	-	-	-
Tx Messages	-	X	X	X	X
RTT, Position Location	X	X	X	-	X
Note: Check indicates more than minimal need for specific JTIDS features					

a. TDMA - Analysis of ET applications indicates that the Phase I TDMA structure is adequate in terms of data capacity and AJ features. As described in Section 4, the TDMA time slots are organized as follows:

- o 12.8 minute epoch consisting of 64 frames
- o Each 12 second frame consists of 1536 time slots
- o Each 7.8125 ms time slot may contain a 109 symbol data message plus sync preamble.

The resulting Lx-band communication system provides a data capacity of approximately 60 K bits/sec.

b. Interoperability - To provide interoperability with existing JTIDS classes of terminals and other ET's, the ET must communicate in the Phase I TDMA waveform. (Alternate "Phase II" waveforms, under consideration by various contractors and the JTIDS Joint Program Office are required to also support the Phase I TDMA implementation). The ET may require use of free text messages for its specialized application. It may process the waveform in a simpler way or not make use of all the AJ features that another terminal can generate. However, the overall ET approach should be that its mode of operation should be transparent to other interrogating or communicating JTIDS terminals.

c. Net synchronization - JTIDS requires synchronization on a time slot by time slot basis. However, it is important to distinguish between system time synchronization and message synchronization. If a terminal solely receives messages or responds to interrogations, and is operating in a minimal AJ mode, message synchronization can be achieved without system time synchronization. This mode of operation is suitable to some of the guided weapon, RPV, and other applications. Where data must be transmitted or received in more complex AJ modes, system time initialization and clock synchronization are required. For short duration missions it may be possible to initialize and maintain sufficiently accurate system time with a free running clock. In other applications, clocks may be corrected via RTT measurements or other periodic message updates, where the accuracy of system time clock synchronization is less than the accuracy required for doing one-way ranging.

d. AJ Modes - JTIDS provides modes which offer various levels of AJ performance by using frequency hopping, PN coding, data interleaving and error correction coding techniques. The minimum level AJ mode is suitable for beacon and sensor applications, and for some guided weapon applications. AJ functional capability impacts the design and cost of a terminal, particularly in the area of receiver design. In addition to the standard operational modes, it may be advantageous

to consider a simpler hybrid ET mode which would give some AJ improvement (this would affect software design of existing terminals). The baseline ET approach should first consider the standard modes.

e. ET Messages - A primary ET requirement is to receive data in the form of commands or RTT interrogations and transmit RTT replies. For mission applications peculiar to the ET, messages may be specially formatted free text JTIDS messages, (although some fixed formatted messages could also be utilized). In addition to receiver design and sync processing, the area of message filtering and processing can influence design complexity, and should be kept minimal. Transmitted messages would also include special purpose free text as well as fixed formatted messages.

f. Position location - JTIDS RTT messages can be used as an aid in position location. As a basic approach, to simplify the ET, location should be determined external to the terminal. If the position can be more conveniently or more properly determined on the ET vehicle (by RTT, TOA one-way ranging or TDOA passive location) because of operational low probability of intercept (LPI) considerations, then the navigation processing may be considered as a modular addition to the terminal, or handled by an external processor.

2.4 EVOLVING TECHNOLOGY IMPACT. In order to meet size, weight, and cost goals for the ET, a number of key technology areas must be identified and exploited. Growth and tailoring to specific applications is achieved by electrical and mechanical modularity and flexible I/O design. Technology elements that impact the design include:

a. Logic design - Use of LSI and microprocessors in the logic design can reduce size of signal and message processors and lower power dissipation which is important in some applications. Microprocessors can aid in performing the special functions, and can be programmed for different applications. Custom LSI, as opposed to general purpose, should be utilized to support those modules or functions which are predictable/common for the various functions.

b. Dense RF packaging - Multiple RF components can be assembled on individual microwave integrated circuit (MIC) dielectric substrates and then interconnected on an overall High-K dielectric board. Besides the obvious size reduction advantages, this approach improves reliability and producibility of the RF hardware.

c. Matched filters - Programmable matched filters are used to correlate the received modulated waveform. For this, SAW devices are more mature in development than CCD devices but CCD's have a projected advantage and should be considered.

d. Transmitters - RF power amplifiers can be a significant cost item in the terminal design. Solid state transmitter designs offer the required power levels of 10 to 50 watts in the JTIDS frequency band at moderate cost with high reliability.

e. Frequency synthesizer - The oscillator design can be a very significant cost item, particularly if frequency hopping is required. Austere, low AJ versions of the terminal could include a frequency oscillator with modular growth for applications requiring frequency agility.

f. Modular design - A number of functions or components are candidates for modularity. In addition to the local oscillator, the power amplifier and message processor may be tailored for different applications.

g. Low-cost packaging - Mechanical structure can compose a significant portion of the weight and cost of the ET. Metalized plastics should be considered for the unit enclosure, to lower weight as well as production costs.

An ET is conceptually depicted in Figure 2-1. The signal/message processor is shown as a single horizontal page. The RF Rx/Tx circuitry is included in a single module (power amplifier, circulator and transmit filter separate). This unit represents the basic minimal requirements ET.

2.5 MECHANICAL AND COST REQUIREMENTS. The size, weight and cost goals for an ET have been previously mentioned. In summary,

- Volume: 300-500 in^3
- Weight: 8-12 lbs
- Cost: \$5K - 10K

As a result of IBM's work on guided weapon and ET programs, these projections are considered to be achievable goals. An estimate of the weight and cost apportionment among the major ET functions is indicated in Figure 2-2. These estimates include the various technology features discussed above, employed to satisfy the basic ET requirements.

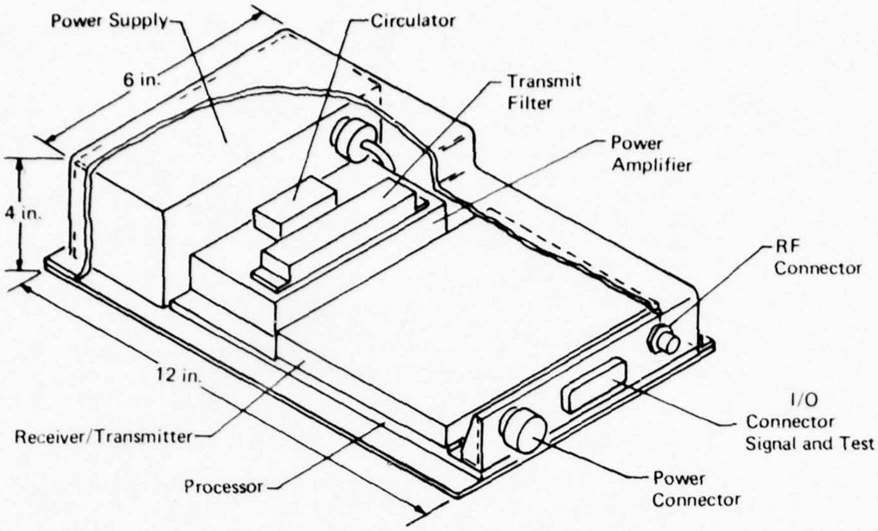


Figure 2-1. Expendable Terminal Mechanical Layout (production concept)

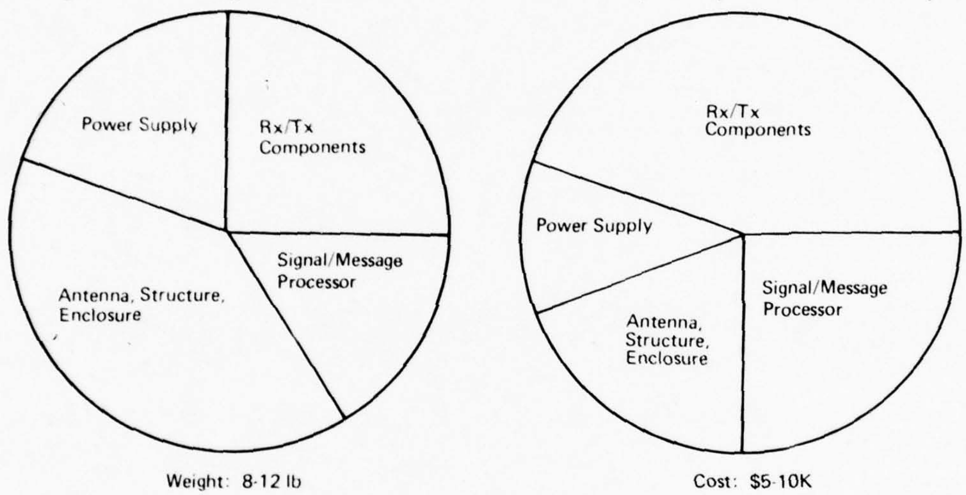


Figure 2-2. ET Weight and Cost Goals

3. MAJOR TERMINAL FUNCTIONS.

An ET must perform four major functions (receive, process, interface, and transmit) to be useful in a variety of applications. These functions are shown in the general purpose ET block diagram, Figure 3-1. The complexity and cost of these functions depend on the required application and which JTIDS features will satisfy system requirements.

There are applications where all four functions are not required. For instance, a terminal to be used for navigation or guidance of a missile may not require the transmit function if it has the on-board processing capability to compute its own position. If the application for the terminal requires it to be fully time synchronized in the JTIDS net, a receiver with programmable SAW or CCD devices would be required which would increase the cost and complexity of the receiver.

3.1 RECEIVE FUNCTION. The receive function of the terminal is to detect and correlate the RF waveform as transmitted by another JTIDS terminal. This spread spectrum waveform includes a range of L-band frequencies which are continuous phase shift modulated (CPSM). After correlation, compressed pulses are sent to the processor.

3.2 PROCESSING FUNCTION. The processing functions of the terminal may be quite diverse and could include a few or all of the following functions:

- Threshold detection on input message
- Control for a frequency synthesizer
- Control for synchronization correlator
- Control for data correlator
- Control for pseudorandom codes
- Symbol interleaving and de-interleaving
- Data encryption and decryption
- Error correction coding and decoding
- Address discrimination and data decode
- Perform TOA measurements
- Perform computations using JTIDS inputs
- Perform computations using input from other sensors

- o Encode data for transmitted messages
- o Provide CPSM modulation patterns
- o Provide control, input, and output for other systems

If the terminal is used for a passive application and is not required to transmit, no processing would be required for data encryption, error correction encoding, symbol interleaving or transmitter modulation.

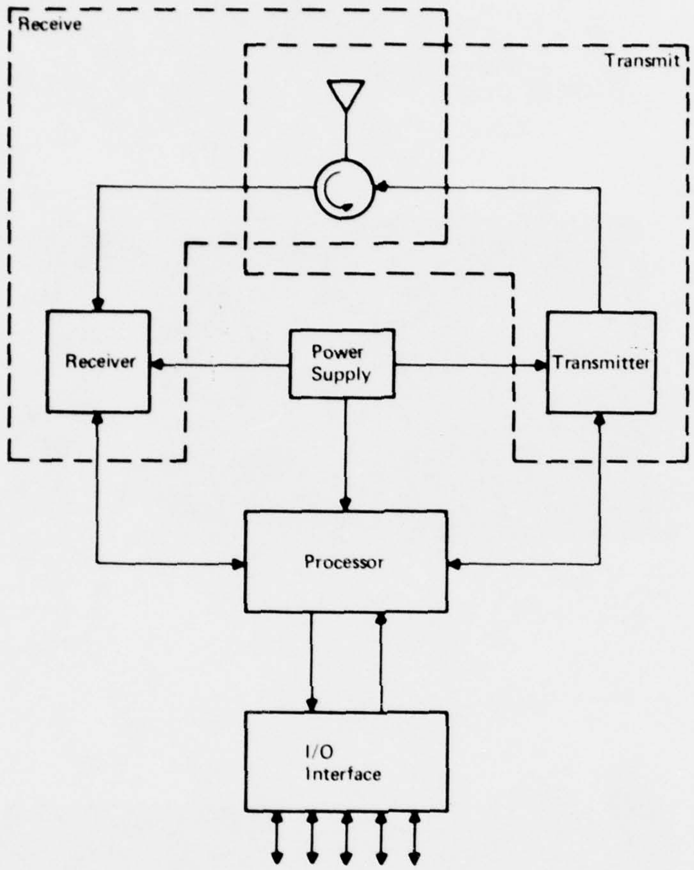


Figure 3-1. General Functional Diagram for a JTIDS Expendable Terminal

3.3 I/O FUNCTIONS. The interface function of the terminal must provide a route for data to and from other systems. Depending on the application, the interface could be used for digital or analog data to or from:

- o Sensors - EW, TV, IR, sonic, pressure, etc.
- o CRT displays - such as an aircraft TV display
- o Analog meters - such as an ADI on missile launching aircraft
- o Other computers - serial or parallel
- o Printers, recorders, or test devices

3.4 TRANSMIT FUNCTION. The transmit function of the terminal is to accept data from the processor and apply modulation to an RF source. This generates the spread spectrum waveform for transmission to other JTIDS terminals. Depending on system range requirements and interrogation rates, a power amplifier would be selected for its output and duty cycle capacity.

4.0 IBM DEVELOPED TERMINAL.

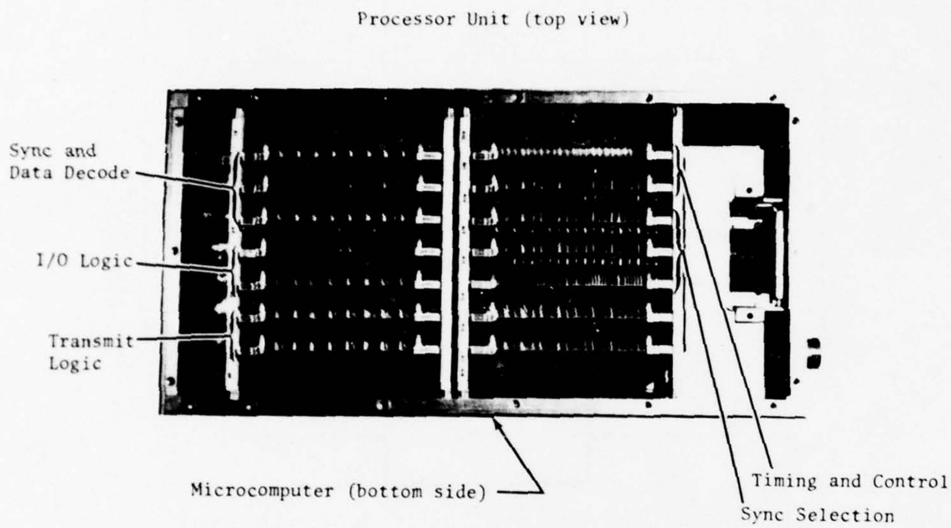
IBM has developed and built an expendable terminal under independent research and development (IRAD) funding which was designed to address cruise missile or guided weapon applications and is compatible with the JTIDS waveform and message structure.

Operationally, Class 2 JTIDS terminals on-board tactical aircraft would send interrogation and guidance messages to the ET. The ET would respond with a transmission which permits the Class 2 terminals to determine the range and location of the missile or weapon. Updated guidance or navigation data, received from a Class 2 terminal, would be provided to the vehicle for guidance and control processing.

The terminal was built as a flyable brassboard and is packaged in two, half-ATR size, units as shown in Figure 4-1. The RF unit contains a receiver, transmitter and power supply while the processor unit contains a microprocessor and digital control - I/O circuits.

4.1 FEATURES. Several features were incorporated into the ET design to achieve a low cost expendable terminal while retaining JTIDS compatibility. Key features of the ET include:

- o JTIDS mode 4 operation
- o Sync on data symbols
- o Single SAW for sync and data
- o (15, 11) Hamming code for error correction
- o Microprocessor for decoder and dynamic calibration
- o Simplified modulation technique



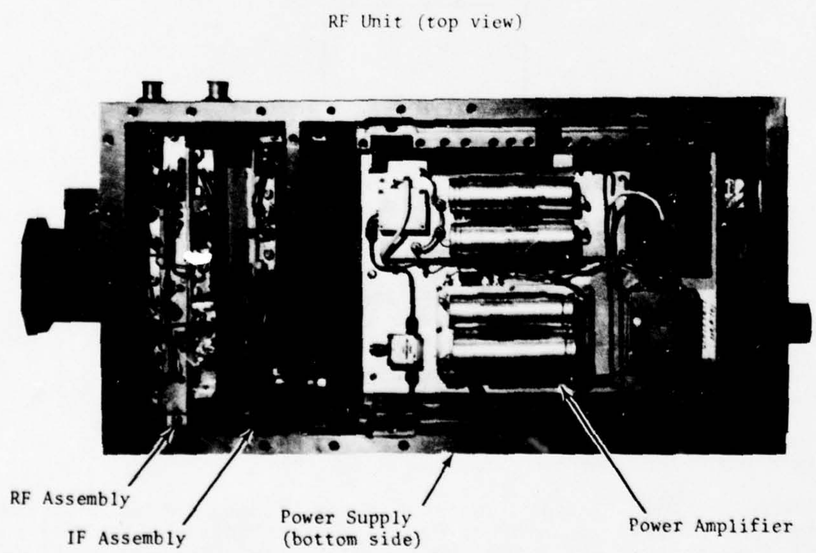


Figure 4-1. IBM-developed JTIDS Terminal (brassboard)

4.2 JTIDS COMPATIBILITY. The terminal accepts fixed frequency, free text messages transmitted in a normal JTIDS net as shown in Figure 4-2. The terminal does not use the synchronization preamble transmitted at the beginning of each time slot and therefore does not require a costly programmable SAW for synchronization.

When an RTT interrogation message is received, the ET uses the 16 cyclic code shift keying (CCSK) header symbols for address discrimination and to establish an accurate time reference for its reply. The header must contain the proper symbols, in the correct sequence, before the ET will respond to it. After a programmable time delay, the ET transmits an RTT reply message which can be received and decoded by the Class 2 terminals. In addition to the normal JTIDS fields (sync preamble, time refine and header), 16 CCSK symbols are transmitted at the end of the message which are used for dynamic calibration of

the ET. The data contained in the header field of the message does not contain TOA information as in a standard JTIDS reply but rather fixed symbol patterns. The phase of the mission profile determines which fixed pattern is transmitted. These patterns are compatible with the (16, 4) Reed-Soloman error correcting code which is required by the Class 2 terminal for proper decoding of message headers.

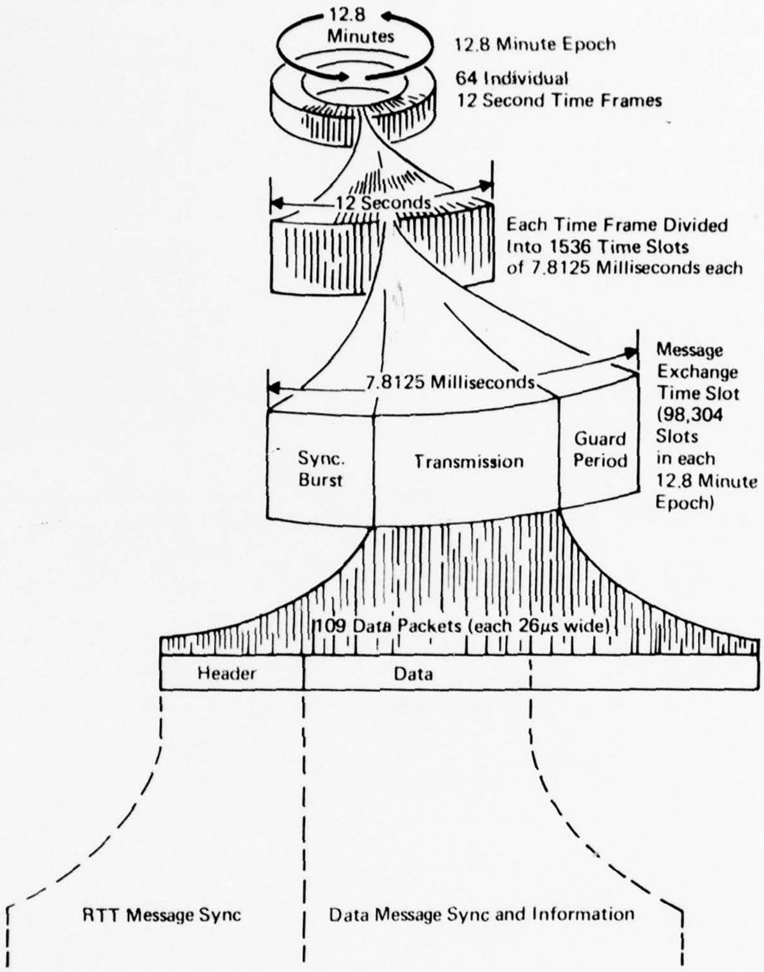


Figure 4-2. ET Use of Standard JTIDS Net

When a type Ø data message is received, the ET utilizes the first 16 symbols of the 93 symbol data field for address discrimination and to establish data decode timing windows. Navigation data is contained in 15 symbols and, after decoding and applying the ECC, 55 serial information bits are provided to the vehicle for implementation.

4.3 FUNCTIONAL DESIGN. A functional block diagram of IBM's brassboard ET is shown in Figure 4-3. The expendable ET presently operates at the single Mode 4 frequency. However, for future expansion, a frequency synthesizer is incorporated in the hardware which is capable of covering a lower portion of the JTIDS band in the required stepping intervals.

a. Receiver - The receiver is capable of correlating the JTIDS waveform from an interrogating Class 2 terminal and will provide a pulse position modulation output to the digital processor. The receiver has 100 dB of blanking during calibration and protection against high power transmissions.

   A SAW matched filter/demodulator (correlator) designed for the JTIDS type modulation is used. A JTIDS message structure was utilized which permitted a SAW to be used for synchronization and data, as a low-cost feature.

b. Processor - The processor uses a threshold detection technique for synchronization. Pulses which are detected at the output of the SAW devices are inputed to a digital delay circuit. If the proper sync address is received, pulses will be correlated; the threshold is set to receive at least 5 pulses to declare coarse sync. A second set of pulses is used for time refinement.

   The time of receipt for each of the fine sync pulses is averaged to obtain the best estimate of message timing.

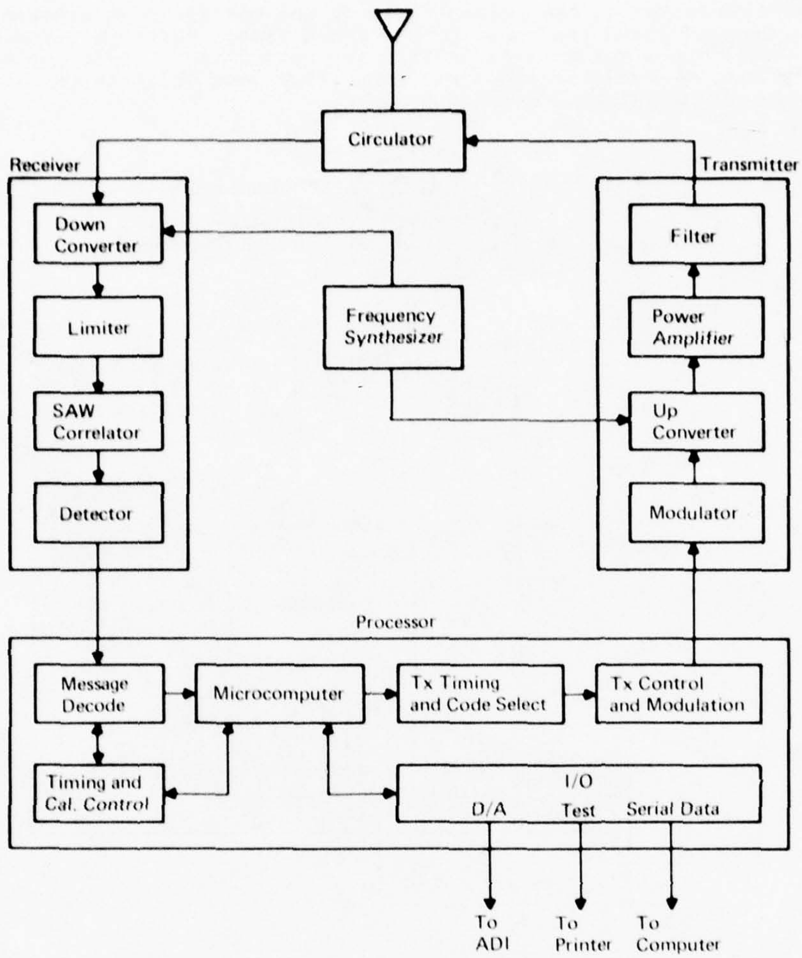


Figure 4-3. ET Functional Block Diagram

The microprocessor is used in error correction of data, control of timing accuracy for dynamic calibration, to provide modulation codes for transmission and to send data to the I/O interface for output to an ADI or digital processor. When the ET receives a DME interrogation it responds after a fixed delay since it is not synchronized to the JTIDS net. This fixed delay must be precise, since the interrogating terminal uses this message to compute round-trip timing for range determination. Equipment delays in the receive and transmit paths which make up a part of the delay can vary with time and temperature. A "wraparound" calibration technique is used where the transmitted signal is "leaked" around through the circulator and received. The wraparound time is measured on each transmitted sequence and is used to subsequently update and correct transmit times so that the overall delay is maintained accurately.

c. Transmitter - The transmitter is capable of transmitting CPSM compatible signals at an output power level of 150 watts at the Mode 4 frequency. CPSM is more expensive to generate than Offset-quad phase shift keying (QPSK) or bi-phase modulation, but has performance parameters that are superior. A modulation technique was developed for the ET which is similar to CPSM in terms of correlation gain and width for the received signal.

5.0 SUMMARY.

There is no intention for ET technology to compete in the arena presently defined for the JTIDS Class 1, 2, and 3 terminals, but merely to augment and expand applications of JTIDS to numerous DoD system requirements. The design philosophy outlined in this paper is grounded upon ET operation in conjunction with and as extensions to "full-up" JTIDS terminals for a meaningful range of applications where low cost, expendability, and unmanned operation are key ingredients.

Responsiveness to these application requirements results in possible design trade-offs in the areas of AJ features, extent of position location/navigation (PLN), degree of synchronization, required clock stability, RF power output, error correction coding, etc. Offloading of certain functions such as PLN and RF gain to controlling JTIDS Class 1, 2, or 3 terminals may be desirable.

These design tradeoffs can be measureably affected by the rapidly evolving logic, processor, and RF component technological progress presently taking place. Low cost/high speed microprocessors and monolithic storage, and SAW/CCD devices are examples of key technology areas.

The IBM-developed JTIDS ET brassboard is an example of cost consciousness, simple design, and practical technology exploitation with sufficient modularity for added functional growth where needed. IBM is working with various DoD agencies pursuing demonstration applications for ET class terminals.

The advent of JTIDS communications and position location holds promise of a revolutionary improvement in combined-arms battlefield coordination giving commanders a level of visibility and control unmatched by present systems. The military services will reap benefits from JTIDS to enhance real-time coordination within and between tactical units, tactical flexibility through mobility, and survivability of the command and control structure despite jamming and battle losses in a hostile environment. The JTIDS ET has a place, along with the Class 1, 2, and 3 terminals, in making JTIDS communications and position location available to the wide spectrum of DoD applications.

INTEGRATION DEVELOPMENTS

Albert WUNSCH
Naval Ocean Systems Center
Code 8152
San Diego, California 92152
U S A

INTRODUCTION

The integration of JTIDS into a platform presents a new twist to the systems designer. For this reason, a number of developments have been initiated to define the nature and scope of the interface required between JTIDS and the system(s) it must service on various platforms. This chapter will touch on several aspects of system integration and describe two efforts presently being pursued by the Navy – The System Lab Test (SLT), and the engineering development model definition (EDMD) program. It is hoped that this will provide the reader with some insight into the integration and application of JTIDS as well as the nature of the integration developments presently underway. As the reader progresses through the chapters, it will become clear that many here-to-fore independent functions will become interrelated and new operational capabilities can be exploited that were not possible before.

JTIDS SYSTEM INTEGRATION

The scope of the integration problem will, of course, vary with the type of platform. It can be a stand alone unit on a small non-NTDS U. S. Navy ship or perhaps a complex integrated terminal on a carrier (CV) or other platform exhibiting heavy communications traffic, many colocated systems, inertial navigation etc. In order to scope the integration task, the JTIDS terminal must be considered from a broad point of view. It must be viewed from the antenna to the man-machine interface as depicted in FIG. 1. This is to ensure that the system operates in the RF world as well as with proper integration of various user system at the baseband level. Referring to FIG. 1, the JTIDS terminal is represented pictorally within the dotted lines. The box between the terminal and the antenna represents those fixes either in this system or other on-board equipments to insure proper isolation between various operating systems. This, of course, must be done without serious degradation to each of the systems involved. To the far right is the user system. These so called users, to site some examples could be voice circuits, the inertial navigation systems (INS), or the (NTDS) Navy Tactical Data System. The interface block represents that which is necessary to allow the user system(s) to access the JTIDS terminal. Before these external blocks can be designed, the systems designer must have an appreciation for the JTIDS terminal. Previous sections have described the JTIDS terminal being developed and it will not be repeated here. However, several points are worth mentioning again to emphasize how JTIDS is different from conventional equipment with which systems designers have had previous experience.

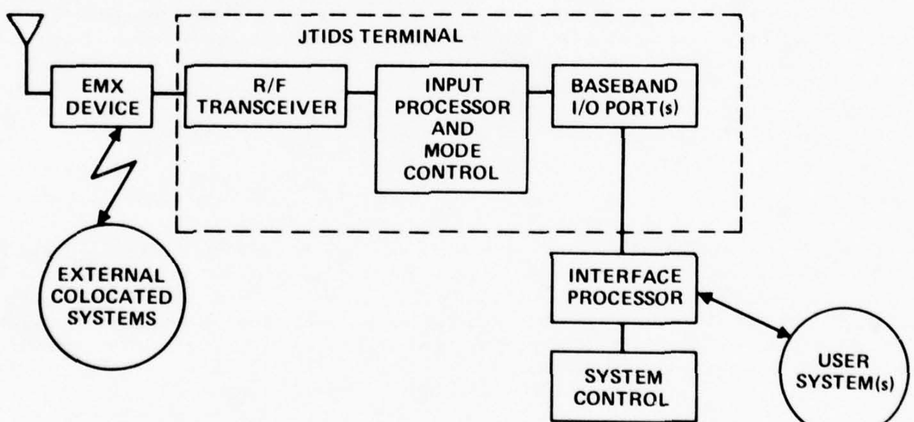


FIGURE 1. JTIDS INTEGRATION SCOPE

The JTIDS terminal is an integrated Communications, Navigation and Identification (ICNI) system. It is felt that this system can significantly impact fleet operations if the key features of "CNI" are exploited through proper planning from the start by the platform developers. Relative navigation (RELNAV) provides, for example, a key element to the type of operation that can be implemented because it permits all JTIDS equipped units to know where they are with respect to other members of the community. The identification feature in conjuction with RELNAV offers new possibilities for positive identification of friendlies. These features are covered in a separate chapter. Communications, on the other hand, is an ICNI feature which is least understood by those who must use it.

JTIDS is not a radio in the conventional sense that most communication system designers and operators are familiar. To the contrary, it is as if there were a number of radios, with varying bandwidth, which must be set up in a way that provides for the most efficient netting of the JTIDS equipped platforms. Consequently, the system designer must consider net management techniques and the determination of the type of service JTIDS is to provide in order to attack the interface design problem. That is to say, the channels to be constructed must be defined in addition to defining the specific interfaces between the various user systems in terms of hand shaking logic, voltage levels, impedance levels, etc. The creation of these channels along with the resultant connectivity between platforms describes a netting scheme with its attendant throughput rates and response times. The designer's task becomes a tradeoff between various channel structure schemes in search of one that is best suited for the application. This is the net management task. It involves not just the individual platform but how the elements of the entire task force are interrelated. It also becomes one of the driving forces to the interface design. It is in this sense that the system integration effort departs from a conventional communications system design task.

To illustrate the notion of defining a channel, consider an example involving fifteen aircraft. They need to establish voice communications and exchange digital data. In addition, a relative navigation message must be sent by each aircraft once every second. FIG. 2, illustrates a TDMA structure with 128 slots per second. For this discussion, assume it takes 3 slots per second for each aircraft to *transmit the digital* data required. Also assume that sixty-three (63) slots are needed to establish the voice capacity. One approach the designer may take is to assign three slots to each aircraft as shown in FIG. 2. The figure shows three per aircraft with slots one through forty-five utilized to provide for all 15 aircraft. Note that separate units of capacity (slots) are assigned to individual aircraft. If one particular aircraft requires more capacity, then additional slots are assigned.

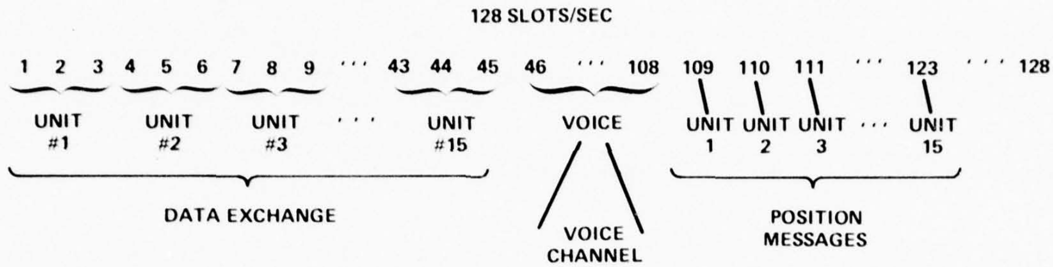


FIGURE 2. ILLUSTRATION OF A POSSIBLE SLOT CONFIGURATION

Another approach could be to establish slots one through forty-five as a single channel. This allows for a faster (15 times greater) exchange of data for a single aircraft, assuming that it can transmit on the next available slot out of the 45, but now requires that a protocol be established among the aircraft to manage this channel. The voice channel has been structured this way. Again referring to FIG. 2, slots forty-six through one hundred and eight (108) establish a voice capacity channel. The interface device buffers this data and switches it to the voice modem for voice communications. The protocol here could be on a first come first serve basis similar to present half duplex voice channels now in use. Slots 109 through 123 are assigned to the relative navigation function on a by-aircraft basis to provide for a position report update once per second. The reader can visualize other configurations and it can be seen that the results affect the nature of the interface.

The data exchange link discussed above could require an interface buffer to store and forward the data from each user to the computer for processing, depending upon the protocol that exists behind the data link. Establishing a functional channel to service a computer system in which certain protocols and message standards are already established is one of the requirements that JTIDS must satisfy. The TADIL-A-link-11 and TADIL-C-link-4A data links presently service the tactical data system in the Navy and represent this type of function. A JTIDS unit can accommodate both of these data link functions by establishing different data rate channels, as illustrated for a different case in the sample above. It is this feature of establishing different data rate channels simultaneously or of interrupting one to construct another on a priority basis that makes JTIDS unique from the standard dedicated data link equipments presently employed. In order to provide the reader with some idea of what is involved in the integration of such a system, the following paragraphs will discuss the interfacing of JTIDS to the tactical data system presently employed in the U. S. Navy.

INTERFACING WITH NTDS

In order to appreciate the implementation, some discussion of NTDS itself and the requirements for the JTIDS interface is needed. The Navy Tactical Data System is an existing information exchange and management system for tactical operations, that is installed on ships, aircraft and ground sites. It is usually referred to as ATDS (Air Tactical Data System) for aircraft, NTDS for shipboard installations, and Marine Tactical Data System (MTDS) for the Marine Corps. FIG. 3, illustrates the major components of the tactical data system. It consists of a data link for transmission of the data, a computer system to process the data in conjunction with the operator commands and a control/display through which operators exercise the system.

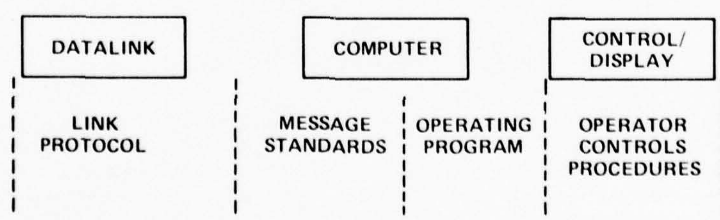


FIGURE 3. TACTICAL DATA SYSTEM (TDS)

It is important that the reader appreciate the relationships of the various elements to each other and the resultant impact they have on the options available to the system integrator. Again referring to FIG. 3, and beginning with the data link function, the link protocol for link-11 on board Navy ships is essentially independent of the computer system and can be replaced. To the computer, the method of transmission is immaterial so long as its requirements are met. A new protocol (the timing and sequence in which participants communicate) can be introduced again, so long as the computer requirements are met. The computer however does place some constraints on the designer because of the design of the operational program and messages employed. The computer requires certain messages in accordance with approved military standards. To digress for the moment, the criticality of message standards becomes clear if one considers the problem of communications between human beings. Several examples will illustrate this need. Consider the exchange of tactical data between two platforms where fuel status is to be reported. In order for the computer system to properly interpret this data, the units of measure must be clear. Is it "pounds of fuel" or "time on station remaining"? Another example concerns the understanding of the command

phrase "engage a target." If this command is given with the meaning "keep track of a particular target" intended but the recipient interprets it to mean "attack" obviously a breakdown of communications results.

The operating program within the computer represents the heart of the TDS and requires certain messages in proper order. If the requirement is not met the input data is rejected. Consequently the system integrator must provide the computer with the standard message set in the order required by the operating program. The requirements can have additional complications as in link-4A where the sequence of the messages from various platforms is significant and the timing of their arrival is critical to the computer and its resultant interpretation. Consequently the computer section places constraints on the adjacent element as to what changes can be accomplished. The TDS can operate so long as the data link provides the transfer of data to the computer in a manner consistent with the operational program – in a sense transparent to the user system. If the design of the operational program is open to change, then a tradeoff can be made between the link/net design and changes to the operational program to obtain the overall performance that best meets the system requirements.

SYSTEMS REQUIREMENTS

Considering an interface which is to be designed in a transparent sense, the system requirements are then defined. JTIDS can provide the data link function for the tactical data system (TDS) by inserting it in place of the existing data link equipment. In order to determine the nature of the interface (i. e. simple replacement vs other) that will satisfy operational needs, the following constraints must be placed on the design. These are:

* No changes to the existing TDS Operational Program (transparency)
* Conventionally equipped units must interoperate with JTIDS equipped units.

The first is primarily a cost consideration. The operations program that must be considered is a configuration that will be in the fleet during JTIDS operational implementation. The second stems from two operational considerations:

(a) Operations involving an over-the-horizon condition where no relay platform is available.

(b) Transition into the fleet.

New equipment is never introduced into the fleet all at once. This results in a mixed force which must continue to carry on fleet operations. Both of the above stated reasons lead to a requirement for a design that can handle both conventional and JTIDS equipped platforms. The approach taken by the Navy is to develop a test bed to verify the compatibility of JTIDS with the Tactical Data System and also to demonstrate the feasibility of interoperating the conventionally equipped platforms with JTIDS equipped platforms. The program charged with this objective is the system lab test. The description that follows will serve to illustrate the nature of the interface design task required to satisfy these requirements. The system lab test will be described followed by a discussion of the implementation being developed.

SYSTEM LAB TEST

The system lab test is a part of the test and evaluation program within JTIDS. The purpose of this effort is to develop the interfaces necessary to test both an integrated JTIDS phase I and phase II system in a simulated operational environment. The test bed consists of the use and interconnection of three Naval facilities on the West Coast – The Fleet Combat Direction Systems Support Activity, (FCDSSA) San Diego, CA; Naval Ocean Systems Center, (NOSC) San Diego, CA and Pacific Missile Test Center (PMTC) at Point Mugu, CA. These facilities will simulate the E-2C, CV and F-14 respectively. One of the objectives of this configuration (in addition to TDS compatibility) is to demonstrate the transmission of M-series messages that are the TACS/TADS Joint Service and NATO approved message standards. The E-2C, CV, and F-14 configurations will utilize the same basic software, written in the CMS-2 high order language on a modular basis and modified (input/output for example) as necessary for each platform to meet platform unique requirements. The software takes advantage of planned compatibility with the AN/UYK-20 and AN/AYK-14 Navy standard computers. The E-2C interface is designed to implement biway (biway is defined in later paragraphs) internal to the E-2C Computer allowing for future data communications growth. The facilities will be interconnected by the use of 50Kbps leased lines which will permit the interchange of tactical data system messages and voice. The three JTIDS units will sit back-to-back with interconnection established at the baseband level.

APPROACH

The system configuration chosen must satisfy the requirements for the interface of JTIDS with the TDS as discussed previously. In the configuration chosen the interface must handle M-series messages in two nets – conventional and JTIDS for what was previously strictly a conventional link-11 net. It must also handle the V & R series messages on link-4A in the same general manner (JTIDS and conventional).

This handling of both allows the interface to affect a cross coupling of the information in both net structures while still satisfying the host computer. FIG. 4, depicts a task force drawn to illustrate the dual operation. The ship designated "N" in the center of the force has the configuration shown in FIG. 5, (as do all the major combatants). To the left of "N" are ships linked together by the current link-11 roll call mode of operation (see FIG. 4a). The other platforms are equipped with JTIDS (FIG. 4b) with the exception of one aircraft. The intent of the design is to provide for an exchange of tactical information between the two groups in such a way that both are actively coordinated in real time operating under a single system protocol. In other words, the platform should interoperate as if there were one, not two, equipments functioning as the data link. This function of two-way exchange between the two groups for the link-11 M-series messages is called "netway."

The aircraft shown at B in FIG. 3b has two aircraft under control. B represents an E-2C with two F-14's linked up for two-way reporting. This linkup represents conventional link-4A or JTIDS transmission of V & R series messages depending upon the aircraft (F-14) configuration. Consequently the E-2C must be able to handle both configurations. In the previous discussion involving the netway, there were two distinct net structures each exchanging tactical information among themselves. The netway here must bridge the gap to allow the entire group (both net participants) to have the same information. In the case of link-4, the information is exchanged between the controller (E-2C or ship) and the aircraft. The problem is one of the controller accessing two different implementations – JTIDS or

conventional. Therefore the problem for the interface is one of routing the commands and responses through either the conventional or JTIDS equipments (as opposed to "netway"-one of information exchange between two net structures). This ability to sort out conventional and JTIDS aircraft messages is called the biway function.

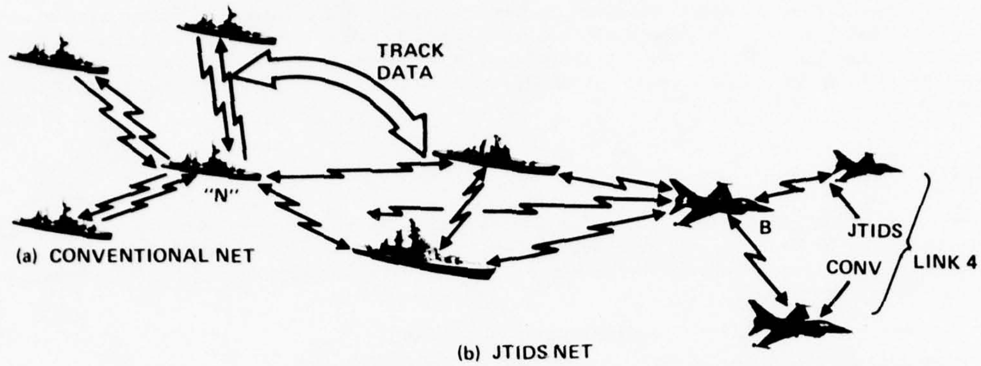


FIGURE 4. TASK FORCE WITH MIXED ELEMENTS

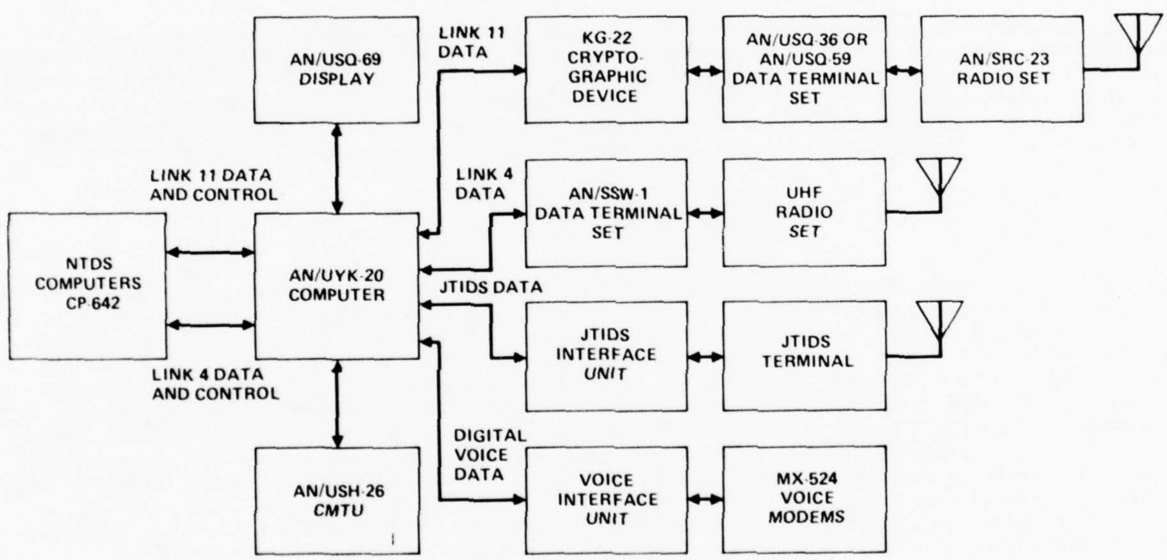


FIGURE 5. NTDS ADVANCED DEVELOPMENT CONFIGURATION

In order to avoid confusion concerning the terminology, some link terms will be defined before proceeding with the interface design discussion. Within the approved military standards are defined the message standards and the data link and link protocol. Thus TADIL-A refers to the "M" series messages and the link-11. TADIL-C refers to the "V & R" series messages and the UHF link-4A data link. In this chapter all references to conventional link-11 will refer to the M-series messages with HF link-11. Likewise link-4 will refer to the transmission of V & R series messages and UHF link-4A. Current link protocols are assumed for conventional links. When the JTIDS link is referred to, the exchange of M-series messages with whatever link protocol is defined will be referred to as the JTIDS link-11 or the link-11 "function." For the JTIDS link-4 case, the reference is to the passing of V & R series messages, again in a JTIDS protocol. The E-2C (B) designation signifies the biway function. It should be noted that the ships also implement the biway function for fighter control or hand-off. Again the need to handle both conventional and JTIDS link 4 is required by the Control platform.

To recap for the moment, we now have defined two groups of link-11 connectivity, interoperable through the "netway," and a need for the biway function in certain ships and aircraft for aircraft control. To facilitate the discussion at this point, let the netway of FIG. 4 also be the net controller (NCS-net control station). Considering FIG. 4 as the example to be satisfied, the approach implemented in the SLT will be discussed to give some insight into what this entails.

SLT SYSTEM CONFIGURATION

The configuration chosen to implement the SLT is shown in FIG. 5. This is an advanced development configuration that will be respecified for engineering development to minimize the computer capacity required. The USQ-69 control/display and USH-26 magnetic

tape unit are AN/UYK-20 peripherals used to initialize and operate the integrated system. The UYK-20 is programmed to handle a number of functions: the packing of messages for transmission through JTIDS; the routing through the conventional route; buffering of received JTIDS messages including unpacking and formating for the NTDS computer etc. The netway and biway functions are both resident in the computer and can be implemented simultaneously and independently of each other. In addition, digitized voice is passed through to the JTIDS units. The MX524 voice modems are being used to provide voice for the test bed.

LINK-11 OPERATION

For the net-way function, three processes occur:

- Conventional link-11
- Relay of JTIDS received data
- Transmitting of ownship data and relay of link-11 received data.

FIG. 6 shows, pictorally, the process referred to. The figure illustrates the transmitted data only for NCS in the center and also for all picket units. (All participants except for the NCS are called picket units.) Not shown is that every picket unit transmission is received by all pickets in the group, as well as by the NCS. The NCS receives and stores for retransmission all received data to be relayed at the proper time. Since the two groups' (JTIDS and conventional) transmissions are neither time synchronous nor of the same protocol, the relative position of the picket response or NCS transmission between the two groups is not synchronized as it may appear in FIG. 6. The JTIDS sequence is for illustration of the concept only. Conceptually, the process shown consists of the UYK-20 allowing the NCS Data Terminal Set (DTS) to roll call the link-11 Picket Units (PU's) in the conventional manner (as in FIG. 4a) followed by ownship information. Ownship information is transmitted on both JTIDS and HF link-11. Following ownship, the NCS relays the JTIDS received data on conventional link-11 to the conventionally equipped PU's. Through its JTIDS unit, the NCS transmits the link-11 PU data to the JTIDS pickets. The forwarding PU procedures are in accordance with operational procedures for relay. Several comments are needed for clarification. As the NCS relays JTIDS received data to the link-11 pickets, it also sends it to its own NTDS computer. The NTDS program on board the JTIDS equipped pickets will not be able to distinguish a change from normal link-11 operations. This is true in the sense that the M-series messages transmitted on JTIDS are packed and unpacked by the UYK-20 and sent to the computer (NTDS) at a rate and in a sequence that satisfies the NTDS computer protocol.

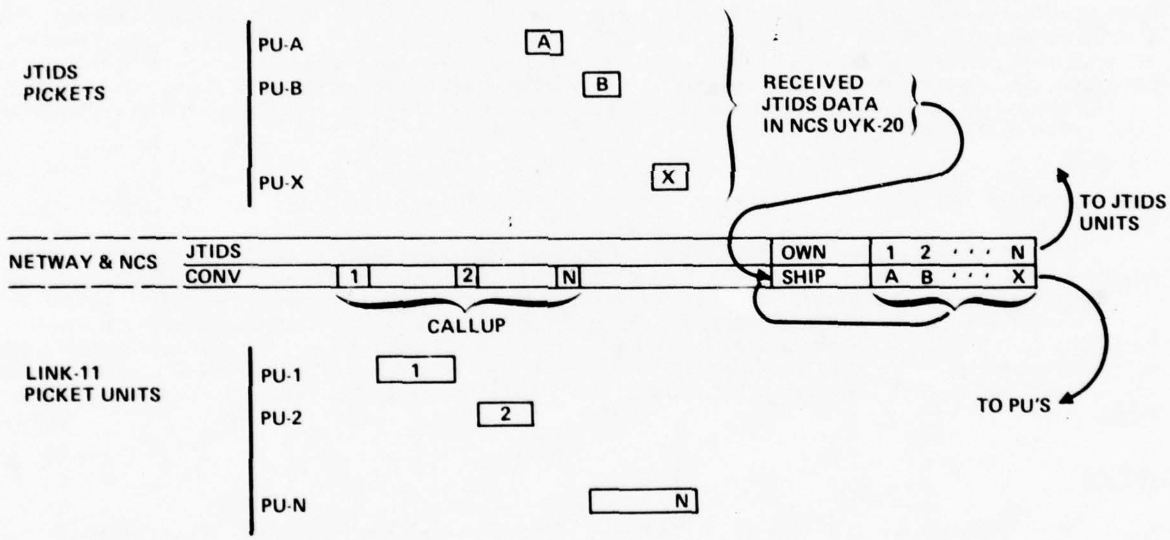


FIGURE 6. FIXED SCOT PROTOCOL ILLUSTRATION

LINK-4 OPERATIONS

For the vectoring and control of fighter aircraft and exchange of track data from the F-14, the operator sets up his aircraft addresses as is currently done. The UYK-20 is told which address refers to JTIDS and which addresses are conventionally equipped aircraft. The conventional protocol has the E-2C (or ship) addressing the aircraft sequentially and receiving the response in sequence. On this basis, the interface needs only to act as a traffic manager. Since JTIDS affords the possibility of increasing both data rates and types of data sent, the system designer has the option of designing a separate asynchronous net. This however, is another problem, outside the scope of this discussion (see IV. B.)

JTIDS INTERFACE UNIT

The purpose of the interface unit is to accept control and message data from the JTIDS terminal in serial format and perform the required data buffering and conversions required to transmit the received JTIDS data to the UYK-20 (see FIG. 7). The reason for this is that the rate of input from the JTIDS does not match the acceptance rate of the UYK-20. Another reason is that of the slot management required to access the unformatted (free test) message port.

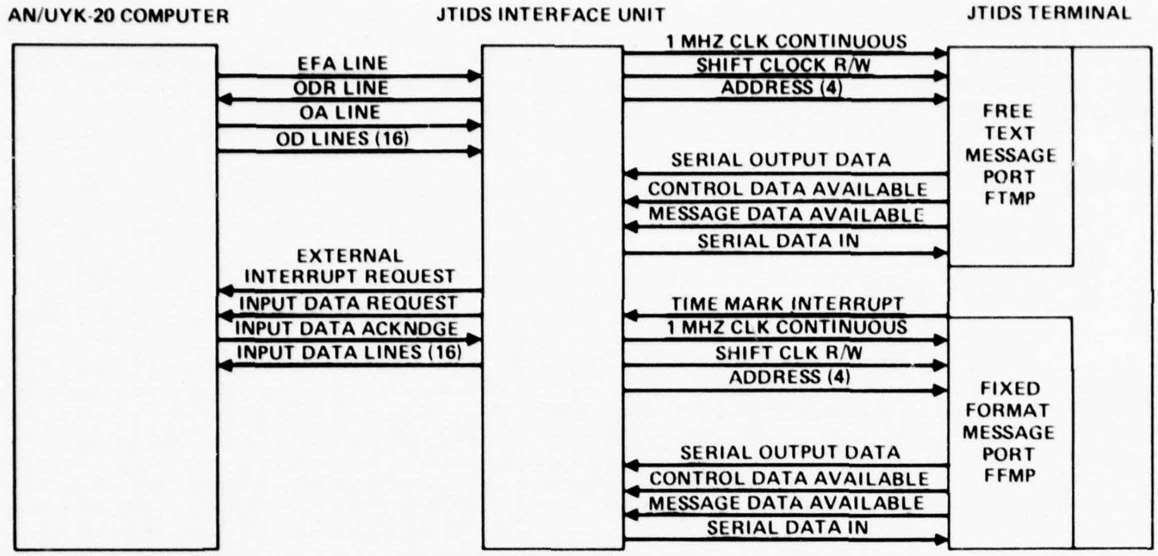


FIGURE 7. COMPUTER/JTIDS TERMINAL INTERFACE COMMUNICATIONS

Previous chapters dealt with the JTIDS unit, it's different message entry ports and the interim message specification (IJMS). The interface in FIG. 7 must be capable of accessing either port in order to operate with units utilizing the Fixed Format Message Port (FFMP) as well as the free text port (FTMP – see FIG. 7) and to initiate the terminal through the USQ-69 Control/Display. This particular interface is designed for the AN/URQ-28 JTIDS terminal (see II.D.(2).a) with the system lab test in mind. The free text port is being used for this integration example for both voice, link-4 and link-11 as defined for the netway and biway concepts. To give the reader some idea of the buffer size, the (JIU) interface for this example would have a maximum capacity of 42 words of 16 bits each (JTIDS to UYK-20) and 61 words UYK-20 to JTIDS.

LINK PROTOCOL

The method of slot assignment bears on the link-11 (JTIDS and conventional combined) net cycle time, response time, access time etc. and impacts the resultant message staleness. The interaction of the number of picket units, their relative data content, the mix of JTIDS to conventional units all affect the aforementioned parameters. Additionally, the content (number of tracks for example) of each participating unit is not the same; consequently a fixed slot assignment for each unit large enough to accommodate a maximum requirement would be grossly inefficient. Simulation of various combinations of picket units and track load distributions provides a guide as to the type of slot structures that can be considered for proper response to the tactical situation of the 1990's. During the SLT testing, several different slot assignment schemes will be investigated. The conventional roll call mode will be one of the configurations demonstrated.

STATUS

The configuration shown in FIG. 5 for the CV as well as a similar configuration for the E-2C and F-14 aircraft test bed will be ready for multi-platform testing in late 1979. The schedule calls for test data to be ready to support a decision process for engineering development in 1980.

The implementation discussed in this chapter dealt with TDMA phase JTIDS. For phase II DTDMA, the system lab test will exercise the same three facilities and although the test phase is not fully defined as of this writing, will exploit the additional flexibility inherent in the DTDMA concept. The phase I tests will exercise link-4, link-11 and voice to verify the ability to operate the netway and biway functions as well as interfacing the tactical data systems.

For phase II testing, TDS compatibility will again be demonstrated. In addition, however, other features will be demonstrated such as voice conferencing and high data rate priority interrupt channels that an Anti-Ship Missile Defense (ASMD) or Over-the-Horizon Targeting (OTH) scenario require. The scope of this simulation will depend upon cost and schedule constraints.

A configuration being evaluated in the system lab test in addition to the E-2C, F-14 and NTDS ship is that of the non-NTDS platform. The JTIDS terminal with a graphic control display will be evaluated as a pseudo picket unit. Since the JTIDS media already has voice and the tactical data message system on the air, a unit configured as in FIG. 8 can receive link-11 data for example. Since all JTIDS platform have relative navigation, a display of all link-11 tracks held in the net can be displayed on a scope relative to the non-NTDS unit. A unit such as shown in FIG. 5 can designate a target to this unit and new tracks can be entered manually into the net through the non-NTDS JTIDS unit.

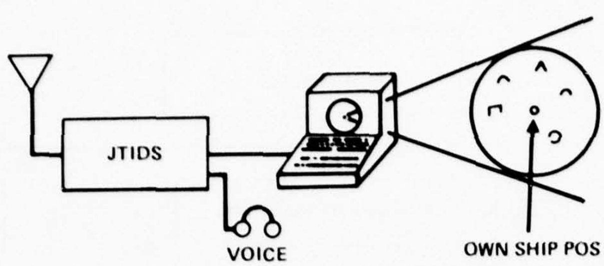


FIGURE 8. NON-NTDS CONFIGURATION

ENGINEERING DEVELOPMENT MODEL (EDM) DEFINITION PROGRAM

The next step in the orderly development of JTIDS is that of engineering development. FIG. 9 illustrates the relationship of the advanced development and related efforts to the decision point (Defense System Acquisition Review Council – DSARC-milestone two) to enter engineering development. The previous sections discussed the system lab test and presented an approach for the integration of a Phase I advanced development JTIDS terminal with NTDS. This effort provides insight into the problems and risks associated with system integration. Nevertheless, it is intended as a test bed and will contribute to the final design and the testing of the final design in the test and evaluation phase of the program.

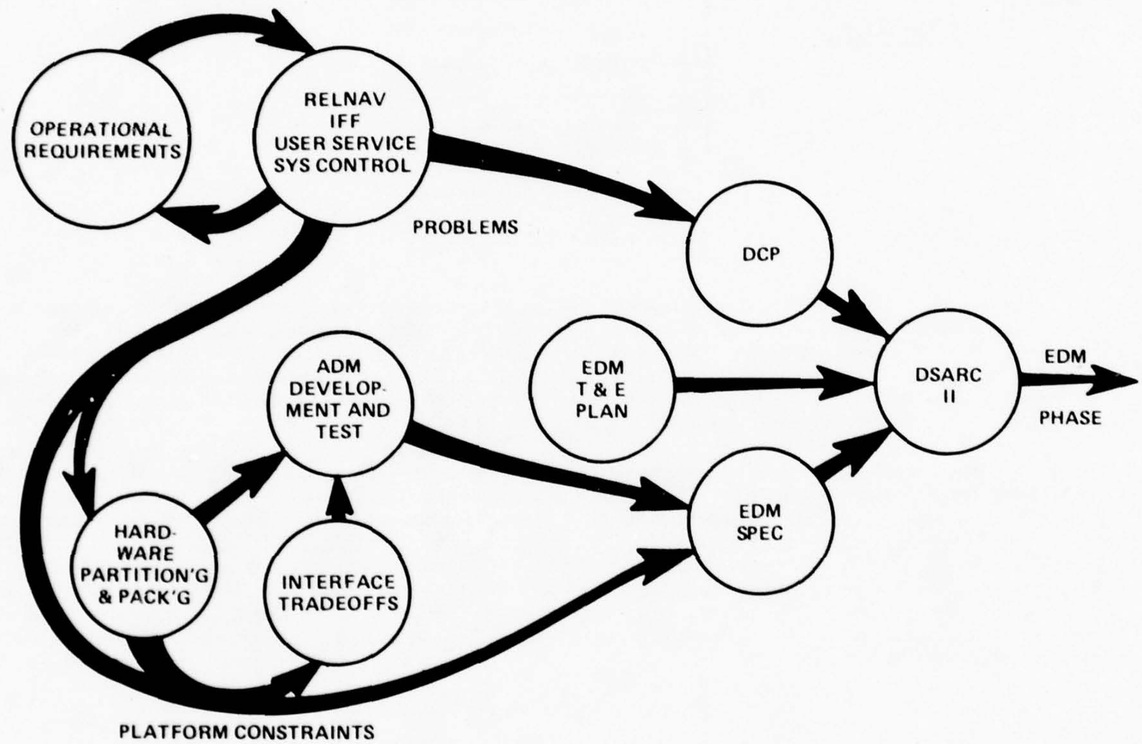


FIGURE 9. RELATIONSHIP OF EFFORTS FOR REVIEW PROCESS

It becomes obvious, however, that an interface, depending upon the type of service required, can be handled by various methods (i.e., microprocessor, minicomputer, etc.). It also appears that this processing could be designed into the JTIDS terminal or considered an external device. This decision, the nature and scope of the interface and its variance of complexity with respect to different platforms must, insofar as they impact the JTIDS design, be determined prior to the definition of the engineering development model. There are a number of efforts underway to describe the operational configuration for ships and aircraft. The EDM definition program is one of these. It is designed to identify shipboard problems and assess the impact these problems will have on the JTIDS terminal design. In addition, it will provide the data necessary to insure an orderly transition onto Navy surface/sub-surface platforms.

PROGRAM DESCRIPTION

As of this writing, the program is just underway. It consists of a number of tests and analysis efforts to define the engineering development model (EDM) of JTIDS for shipboard use as well as the functional configuration, for various surface/subsurface platforms.

The areas of investigation as shown in FIG. 10, consist of three general areas. The RF environment deals with the outside (topside) platform environment. The specification area deals with the internal platform oriented problems and the actual generation of the specification; and the third area deals with a number of tasks necessary to define the system, its requirements and alternate solutions.

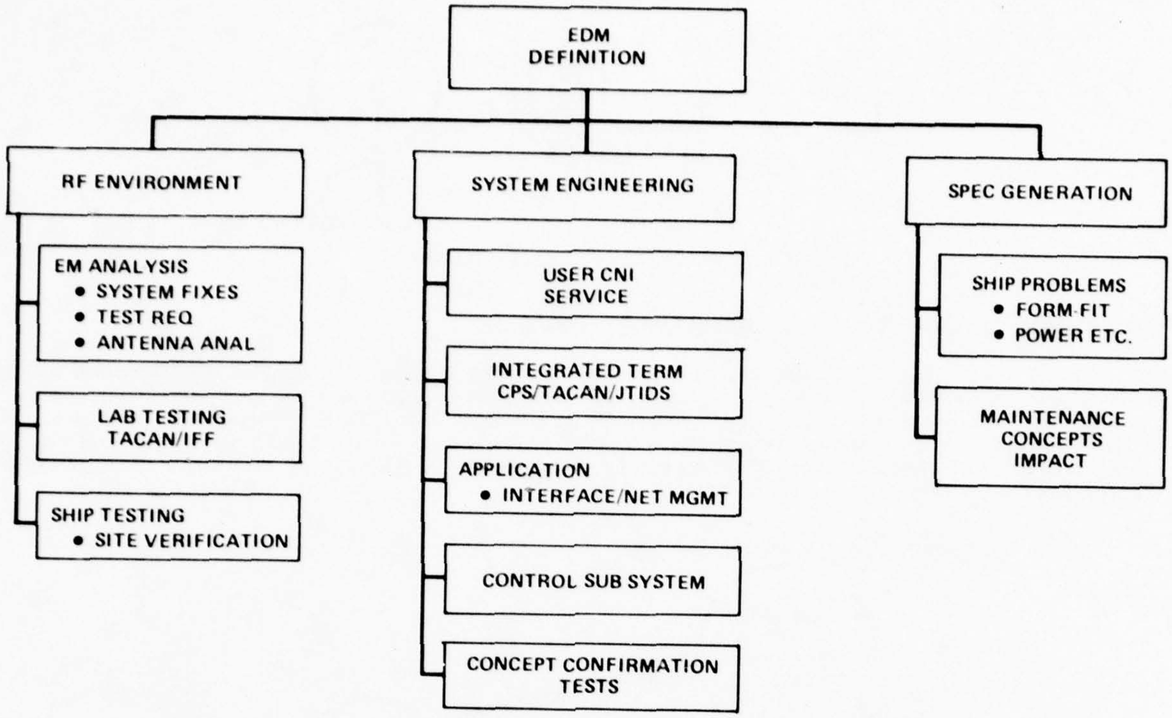


FIGURE 10. EDMD AREAS OF INVESTIGATION

The approach is to define the type of service, the interplatform connectivity requirements and the platform user systems, in order to establish the base from which the interface design tradeoff analysis can be conducted. In a parallel effort, the impact of environmental requirements on the JTIDS design as presently configured (e.g. co-located TACAN beacon) will be assessed. From these analyses, surface/subsurface needs can be identified and the platform configuration of JTIDS, including interface antenna and man-machine interfaces can be determined and evaluated. This allows for a logical development of both JTIDS and those modifications that may be required of the user platforms with the intent to minimize the transition impact on naval operations.

From an architectural point of view, the net management concept (how interplatform information is handled – channel/net protocols) and user service (how much of what type of information) impact most heavily on the interface design, the installation on a platform and the impact of the type of terminal (Phase I or Phase II) that is selected. The objective of this program is to identify the requirements, design alternatives, surface/subsurface needs and from this data to define the surface/subsurface EDM design(s) using the existing JTIDS design as a springboard. The objective is to maintain commonality across platforms to the maximum extent possible. The program has been structured to provide the necessary data in time for the DSARC II decision process. Additionally, in conjunction with other programs, critical items will have been identified and test data made available to demonstrate feasibility where applicable. User needs and installation problems will have been addressed prior to the manufacture of engineering development models in order to minimize life cycle costs and increase the probability of meeting unit cost.

SUMMARY

The intent of this chapter is to impress upon the reader the different nature of JTIDS and a need for a proper integration if JTIDS is to be exploited to the fullest. The SLT approach provides some insight into the nature of an integration with the tactical data system. Those familiar with the tactical data system can appreciate the potential improvement in C^3 possible through the exploitation of JTIDS by future TDS systems. The two-pronged approach (SLT in parallel with the FDMD) to prepare for engineering development is to insure a sound engineering design. (Critical items will have been identified and test data made available to demonstrate feasibility in critical areas. User needs and installation problem will have been addressed prior to the manufacture of engineering development models which should keep the total life cycle costs down and increase the probability of meeting unit cost goals.) Other programs for aircraft and ground site integration which cannot be addressed in this chapter are also underway in order to provide all the data necessary to facilitate the decision process toward a final design. Details on the SLT implementation can be found in documents listed in the bibliography following this chapter. It is hoped that this chapter will help the reader to understand the utilization and application of JTIDS from either a designer's or operational user's point of view.

REFERENCES

1. Naval Ocean System Center, "Benefits to AEGIS Offered by JTIDS – A High Capacity CNS System," T. Shapard and R. Kelly, Sept 30, 1977, Confidential Report

2. Naval Electronics Laboratory Center, "Link-11 With Subnets: An Analysis and Simulation," E. Pasahow, June 16, 1975, Confidential Report

3. Naval Electronics Laboratory Center, "A Study of JTIDS Link-11 and Link-4 Communications Networks," E. Pasahow and E. Hayes, Dec 1, 1976, Confidential Report

4. Naval Electronics Laboratory Center, "A Simulation and Analysis of JTIDS and Link-11 Communications Networks With a Gateway Unit," E. Pasahow and E. Hayes, Sept 7, 1976, Confidential Report

5. Naval Ocean System Center (NOSC), "System Lab Test Specification for JTIDS Shipboard Integration System," E. Pasahow and E. Hayes, May 27, 1977, Confidential Report

6. NOSC, "Preliminary Interface Design Specification for JTIDS," July 17, 1977, Confidential Report

7. NOSC, "System Segment Specification for System Lab Test," W. Sauer, Oct 4, 1977, Unclassified

8. NOSC, "Theory of Operation – Voice Interface Unit," S. Morton, Oct 6, 1977, Unclassified

9. NOSC, "JTIDS Interface Unit for System Lab Test," H. Gerard, Dec 1977, Unclassified

DISTRIBUTED TDMA
AN APPROACH TO JTIDS PHASE II

BY

J. RUBIN
ITT AVIONICS DIVISION
CNI SYSTEMS DEPARTMENT
500 WASHINGTON AVENUE
NUTLEY, N.J. 07110

SUMMARY

The Joint Tactical Information Distribution System (JTIDS) is a tri-service, multi-channel, multi-function system. During Phase I, a new Time Division Multiple Access (TDMA) channel (JTIDS I) was introduced to handle total connectivity information distribution for digital communications and relative navigation. Distributed TDMA, an approach to JTIDS Phase II (JTIDS II) expands this capability by providing increased data rate capability and additional C-N-I function. DTDMA also develops a channel architecture to permit a single terminal to participate concurrently in multiple independent functions or network which have the flexibility to be structured and organized in a manner to efficiently meet the broad spectrum of tactical information-distribution operational requirements. JTIDS II DTDMA is compatible with and includes the TDMA function of JTIDS I which becomes one of the channel structures available to system users.

ITT Avionics Division is currently developing, under JTIDS Joint Program Office/Naval Air Development Center contracts, Command/Control and Tactical Fighter Terminals with the JTIDS II/DTDMA architecture. These current ITT Avionics developments of the JTIDS Phase II/DTDMA system have a firm technical base established by previous program developments in the area of CNI multi-purpose systems. These related developments operated at Lx-Band; established the low-duty signal structures and the Reed-Solomon code; performed communications-navigation-identification functions concurrently; and utilized new advanced processing techniques and algorithms which are directly applicable to JTIDS II/DTDMA.

JTIDS II/DTDMA is a proposed solution to meeting the broad range ICNI requirements for next generation tactical command-control-communications, precision navigation and positive identification.

The utility of this DTDMA technique is shown in its application to the JTIDS II/DTDMA requirements for a C^3/ICNI system. This is highlighted in the common information medium (band/channel) usage of a large number of "Statistically Orthogonal" *time-frequency-phase code cell patterns. It will be shown that a large number of C^3/ICNI users/functions can simultaneously and independently access the information channel with an acceptably low probability of mutual interference. The signals developed via this DTDMA technique are noiselike in their distribution and any mutual interference is readily accommodated by well developed error protection techniques.

1. INTRODUCTION

1.1 BACKGROUND

The necessity for major improvements in tactical aeronautical command-control-communications (C^3), navigation, and identification has become apparent. Military aircraft and related weapon systems can no longer be depended upon to perform effectively in the current and postulated threat environment without improvement in the quality of their CNI systems. Secure, anti-jamming digital communications are essential to efficient command and control. In addition, flexible highly accurate relative navigation is required to complement the present geodetic and self-contained systems to satisfy the multi-mission requirements of the highly mobile tactical environment.

When these requirements are examined in the light of funding and spectrum constraints and with the realization that the use of existing CNI systems will continue for at least another decade, the problems of developing and implementing a new system are complex.

To appreciate the added command and control capability that a distributed communication link (such as DTDMA/JTIDS II) offers the tactical military community, a brief review of operational tactics is required.

The nature of the multi-platform command and control communications (C^3) problem involves real-time coordination of many diverse user terminals. Each of these participating

*Mutually exclusive on a statistical basis.

terminals has specific functions which relate to specific C^3 requirements. It is clear that a large central command platform such as an aircraft carrier has substantially different needs than a tactical fighter terminal or an expendable missile terminal.

To date, conventional communication system designs which address the C^3 problem in light of these diverse requirements provide a compromise solution. This compromise usually favors the more sophisticated terminal requirements and penalizes the tactical and reduced requirement users with an unnecessary system capability and hence excessive expense. In JTIDS II/DTDMA the performance can be tailored to specific needs, thus resulting in a more consistent cost/requirement relationship.

Many of the platforms involved in typical tactical communications or mission scenarios have very specific connectivity relationships. The ability to operate as groups of terminals on specific functionally connected nets or communities is highly desirable and efficient.

Figure 1 shows the JTIDS II/DTDMA Integrated CNI operational utilization concepts. Both large command terminals and austere manpack terminals are included. The full complement of COMM-NAV-IDENT is provided. From Figure 1 we highlight the basic C^3 capabilities for secure and anti-jamming digital data and digital voice. Precision relative navigation and conventional TACAN navigation is included. A combination of the digital data and precision navigation is utilized for coordinated EW operations.

To meet the expanding scope of modern tactical C^3 requirements, the DoD has currently under development, as part of a tri-service C^3/ICNI system, the Joint Tactical Information Distribution System (JTIDS). The initial JTIDS implementation developed a TDMA approach. The second phase, JTIDS II, is developing a Distributed TDMA as a candidate solution.

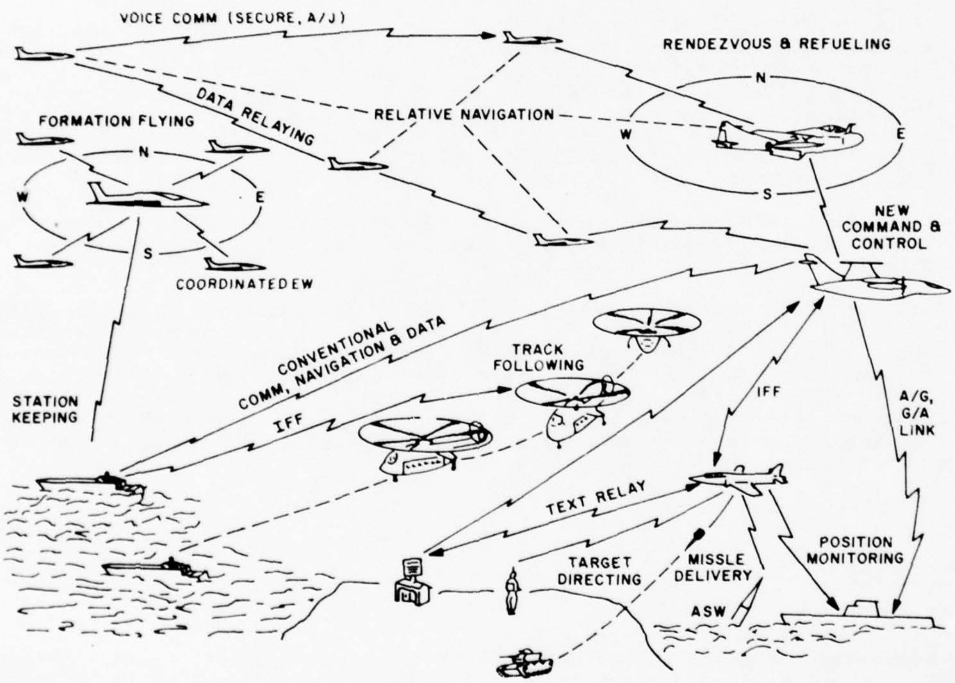


Figure 1. JTIDS II/DTDMA - Integrated CNI System Operational Concept

1.2 DISTRIBUTED TDMA OVERVIEW

Distributed Time Division Multiple Access is a technique that employs low duty cycle pulses, pseudo-randomly distributed in the time-frequency-phase code domain. This T-F-C ambiguity is maintained for all the C^3/ICNI functions on a pulse-by-pulse basis, thereby greatly enhancing random user/function access.

DTDMA signal structure flexibility permits a system architecture which can be tailored to the composition of the particular communications being sent. Rigid time separation is not required, so messages are not constrained to any slot of fixed time duration. Consequently, a message uses only that portion of the system capacity which is actually required, permitting system resources to be used and reallocated with maximum efficiency.

This approach also permits the system, its channels and its terminals to be configured with soft capacity limits. The statistical nature of distributed signal structures is such

that the consequences of responding to peak capacity demands do not cause service interruption. Only momentary reduction in anti-jamming margins result.

This efficient utilization of terminal resources is especially significant for many of the more austere platforms involved in typical tactical communications scenarios. In many scenarios, simultaneous participation by terminals in several independent nets is required to meet total mission connectivity demands. The different types of tactical information which must be exchanged vary in message length, message rate, access type, housekeeping needs, coding requirements and message construction. Because of pseudo-random T-F-C separations and the flexible channel design available through the use of the new DTDMA system architecture, the multiple independent time netting and variable information parameters can be readily achieved.

The feasibility, multi-function and multi-platform applicability, high performance anti-jamming secure communications make Distributed TDMA a particularly attractive candidate for JTIDS Phase II.

2. SYSTEM DESCRIPTION

JTIDS II/DTDMA operates in Lx-Band (960 to 1215 MHz) (See III D). It provides a system that combines spread-spectrum, multiple-access command and control functions, with conventional Tacan navigation and IFF identification service. The system uses the full ambiguities of time-frequency-code to provide a digital communication service that offers high levels of anti-jamming and low probability of exploitation coupled with powerful pseudo-noise access techniques which make real-time C^3, anti-jamming (A/J) and low probability of exploitation (LPE) a practical reality.

The following paragraphs present a brief overview of the JTIDS II/DTDMA system, with a review of several important system parameters.

2.1 CHANNEL ARCHITECTURE

JTIDS II/DTDMA channelization and signal structure are based upon maximizing the utility of available ambiguities in time, frequency and code. This multi-dimensional approach simultaneously contributes to efficient multiple access (MA), anti-jam (A/J) and low probability of exploitation (LPE) system solutions. Moreover, it permits co-existence with current users of the band without significant performance degradation to either the old, or potential new, band occupants. The multiple access problem is solved by a pseudo-random, time/frequency (T/F) hopping, interleaved channelization, utilizing low duty signal structure. Random access low duty techniques provide excellent intra-system and co-band interference rejection, and permit the sharing of transmitters, receivers and signal processors for the various communication, navigation and identification (CNI) functions. The A/J problem is simultaneously solved by the hybrid combination of the above T/F hopping modulation plus the coded, spread-spectrum techniques applied to the signal pulses. A desired by-product of this approach to the A/J problem is the associated high resolution and accuracy in the measurement of time-of-arrival (TOA) of the signal pulses due to increased pulse modulation bandwidth. This precision ranging capability makes accurate Relative Navigation practical. (See III B). Low Probability of Exploitation (LPE) is also enhanced with the use of time and frequency ambiguities as much of the signal predictability is removed by the signal and channel design.

The pseudo-random time-frequency structure results in a "Statistical" (as opposed to rigidly deterministic) orthogonality. This statistical orthogonality designed into the system provides not only the above solutions but also minimizes the need for rigid organizational aspects which are normally associated with other types of multiple access system approaches. Moreover, multiple independent nets of users are capable of sharing the band and individual users may currently participate in multiple nets without major changes in the terminal design.

If all of the Lx-band time/frequency (T/F) space is subdivided in elemental T/F resolution elements, they form T/F patterns which may be identified with RF transmission channels for use by selected terminals for specific functions (Figure 2). These channels are, in general, classified and distinguished by channel data rate and the type of channel access afforded. The DTDMA system architecture provides flexibility in channel creation and can avoid the pitfalls of creating high-capacity, under-utilized, rigid channels. The total Lx T/F band is shared in common by large numbers of terminals, for many functions, utilizing channels tailored specifically, and capable of being programmed dynamically to current requirements. In conventional high duty, or continuous transmission modes, the message synchronization preamble is followed by the data message for each user in turn. In DTDMA both synchronization and message signals, from many terminals, are dispersed in time (and frequency), resulting in interleaved or parallel transmissions from the many sources. Where these distributed techniques are employed, there is no high concentration, with their corresponding vulnerabilities to intelligent interference and intercept.

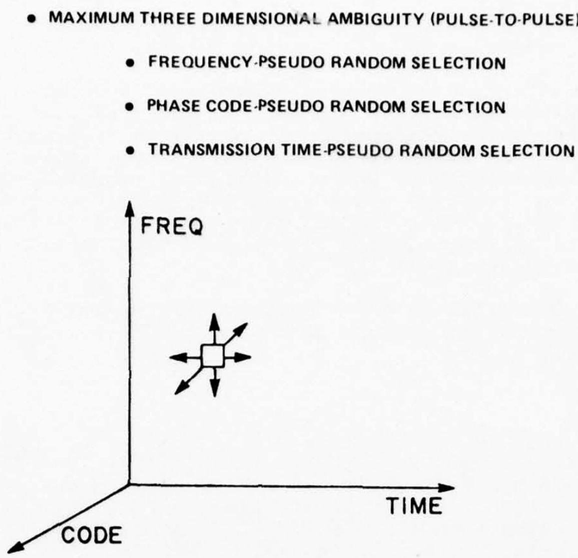


Figure 2. Time-Frequency-Code Ambiguity

2.1.1 Basic Event Interval

Each channel design for DTDMA employs pulse patterns derived from a sequence (or system) timing source producing 78,125 contiguous events per second. Each event interval (12.8 microseconds) is unique and is associated with a specific discrete value of sequence time and a corresponding pseudo-randomly (PR) generated event (binary) codeword.

Assigned bits of this PR code word for each of the 12.8 μs intervals or events (78,125 per second) will uniquely define all signal and channel parameters.

The following key channel parameters are determined from the PR code words for each event:

- Transmission Event
- Reception Event
- Unused Event

If the event is scheduled for either transmission or reception, the following additional parameters are extracted from the PR code word:

- PR Transmission/Reception Time
- RF Carrier Frequency
- PR Chip Modulation Code
- Data Event or Synchronization Event
- PR Data Scrambling

Figure 3 presents a simplified example of PR event code usage.

If all terminals operating on a particular net are at the same discrete value of time, the same codeword is generated at all terminals at this time. Multiple nets are created when several sequence timing chains (or net times) are offset in real time from each other. For a terminal to participate in multiple nets, these offsets from its own (terminal) clock must be determined. By computing these time offsets, a terminal is able to develop replicas of other net times and thereby, generate the appropriate event codes associated with these nets.

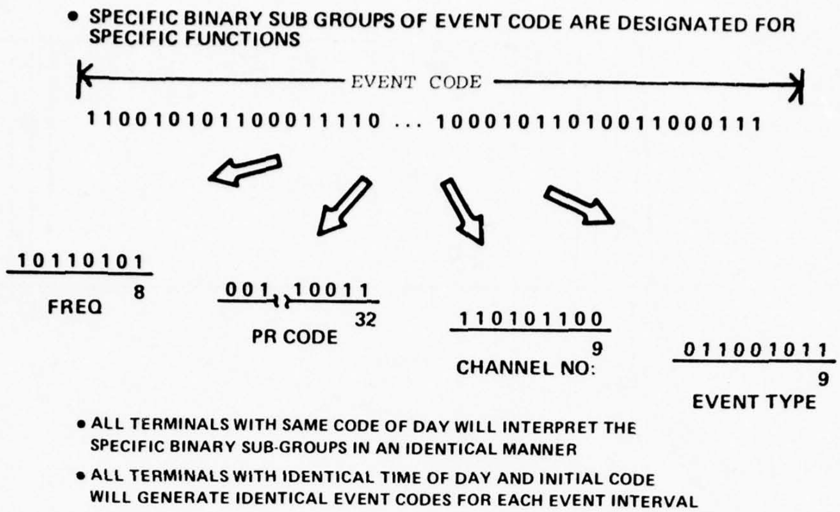


Figure 3. Pseudo-Random Event Code Usage

2.2 GENERIC SIGNAL TYPES

Before proceeding with a more detailed discussion of the DTDMA Modular Channel concept, some definition of generic signal types is required.

Signals transmitted on various types of JTIDS II/DTDMA channels are selected from a generic set of signal types. They include various kinds of synchronization signals, net entry signals and data messages. Before describing these signals, it is helpful to describe the specific JTIDS II/DTDMA pulse structure basic to all of these signals.

The principle unifying element that makes compatible integration of C, N, and I functions feasible is the low duty pulse signal structure. This choice comes from ITT Avionics' broad background of experience with pseudo-random low duty signals which are basic to Tacan, IFF, Air Force Cooperative Countermeasures Link and the Navy ITACS Demonstration systems. Whereas the Tacan (Gaussian) and IFF (rectangular) pulse shapes and pulse durations are determined by existing specifications for each, the pulse structure for the other CNI functions (A/J Communications and Precision Navigation) had to be selected. After considerable discussion among representatives from the armed services and industry, a spread spectrum, PR-coded, 6.4-microsecond duration rectangular pulse was selected. This choice was based on trade-off considerations between large A/J processing gain, transmitter duty factor/peak power capabilities, co-band intrasystem and intersystem interference rejection capabilities, and denial of frequency tracking jamming strategies. Therefore, the JTIDS II/DTDMA signal structure consists of three types, as follows:

Type I:	Gaussian, 3.5 microsecond (Tacan)
Type II:	Rectangular (IFF) 0.5/0.8 microsecond, Interrogations 0.45 microsecond, Replies
Type III:	Rectangular (A/J comm, Precision Nav), Spread Spectrum, Coded, 6.4 microseconds

Since Type I and Type II are well defined, emphasis will be provided on describing the Type III pulse.

The basic Type III (Figure 4) pulse is of 6.4 microsecond duration, spread in spectrum by a phase modulation code of 32 chips, with each chip of 200-nanosecond duration. The carrier frequency is pseudo-randomly selected from 150 frequencies, within the Lx-Band of 960-1215 MHz, and hopped from pulse-to-pulse. Guard bands about the IFF frequencies of 1030 and 1090 MHz are excluded as are the guard bands at the upper and lower band edges. These pulses are used for the functions of synchronization, ranging and data transfer.

The PR modulation codes for synchronization and ranging are pseudo-randomly selected and vary from a random selection of 32 chips, to a specific chip sequence selection from a set of codes with good one-shot auto-correlation (sidelobe) properties. The choice depends on the mode of synchronization (coarse or fine) and the A/J and spoofing protection required. The PR code for data is one from the set of good cyclic code shift keyed (CCSK) sequences, characterized by low cross-correlation between all of its cyclic shifts. The data code is one of the 31 chip maximal length sequences augmented by an additional chip and is identical to the sequence used in TDMA (See IVA).

The JTIDS II signal structure may be characterized as belonging to the class of frequency-hopped, time-hopped, pseudo-noise, multiple-access waveforms.

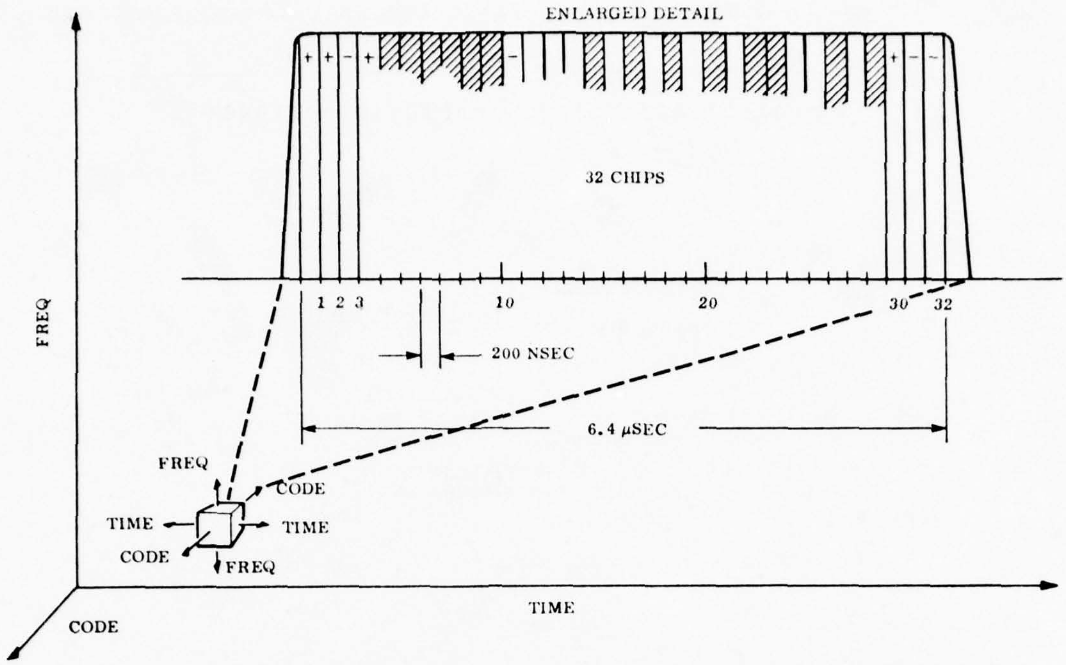


Figure 4. Basic JTIDS/DTDMA Type III Pulse

2.2.1 User Source Synchronization Signal (C)

User source synchronization signals designated as (C) are Type III pulses associated with the Basic Events of a unique signal Basic Low Event Rate Subchannel (BLS) which is assigned to each terminal of a net for its transmissions only. User source synchronization is used by receiving terminals to maintain tracking of the expected time of arrival for each net member. The C pulses are generated at an average rate of about 5 Hz and occur pseudorandomly in time. The RF frequency and the phase code of each C pulse are selected pseudorandomly from the bits of the PR codeword associated with the Basic Event.

2.2.2 Message Start Signal (MST)

The message start signal designated as (MST) is part of a synchronization preamble and is used to signify by its detection that a message is to follow. It is always associated with a channel access opportunity, i.e., a Message Start (M), and is a sequence of Type III pulses which are generated and transmitted at a rate determined by a Basic or Composite Channel event rate for a specific channel design. The parameters of these Type III pulses are similar to that of the C pulses and, thus, the receive timing accuracy is maintained.

The MST signal is, in general, made up of two parts: Message Start Signal-A (MSA) and Message Start Signal-B (MSB). Either MSA and MSB are both used or only MSB is required. For closed community channels, in which all users are tracking each other's C pulses and thereby maintaining an accurate measure of expected time-of-arrival (TOA), the shortened message start signal consisting of only MSB is used. For open community channels, without user source synchronization tracking where no expected TOA is available, the full MST signal consisting of both MSA and MSB is used to enable a receiving terminal to search out the range-time uncertainty between transmitting and receiving terminals.

2.2.3 Message Fine Synchronization Signal (F)

The message fine synchronization signal (F) when transmitted following the MST signal, forms the second part of the synchronization preamble. When transmitted alone, without the MST signal, it is the total synchronization preamble. The F signal is also associated with a Message Start M and is transmitted at the channel event rate of a Basic or Composite Channel for a specific design. The F signal is used through a tracking loop and averaging process to improve the received pulse timing accuracy to a level which is required to demodulate data.

The F signal is comprised of two parts: Fine Synchronization Signal-A (FSA) and Fine Synchronization Signal-B (FSB). The FSA signal is used to improve the time-of-arrival (TOA) synchronization accuracy of the received pulses, and to reduce the probability of a false alarm triggered by noise. The FSB signals are used by the tracking loop to further refine TOA accuracy through additional averaging.

2.2.4 Message Source Synchronization Signal (S)

For messages of long duration, (several seconds), an additional synchronization signal is provided to maintain chip synchronization. The message source synchronization signal (S) provides this function and is interleaved with data pulses at some low rate. It is convenient to designate the rate as some multiple of the Basic Low Event Rate Subchannel (5 Hz).

2.2.5 Header Word (H)

For some messages a header word (H) is added to the synchronization preamble (MST and F) prior to transmission of digital voice or data words. The header word is a Reed-Solomon (R-S) codeword (see paragraph 2.2.7) with up to 14 characters (70 bits) available for information, such as User Identification Number, Message Label, Priority Assignment, Subchannel Address, User Address, Time Quality, Outer Parity Checks on other RS Codewords, etc. It can be also used as a steering or signaling word to direct the receiver to one of several possible subchannels for subsequent message transmission.

2.2.6 Date Message - Uncoded

For a high performance information distribution system, in benign or hostile environments, it is desirable to provide reliable and flexible communication with a choice of both uncoded and coded transmissions. A linked doubly-coded (concatenated) channel is a preferred scheme (for a given code length) since decoding is performed on two or more simpler codes and the overall decoding complexity is reduced. Figure 7 shows a block diagram of a two level concatenated coded channel.

For uncoded messages, the data link performance is determined by the inner channel. The inner channel is M-ary (M=32) in keeping with the desirability for a low duty signal structure.

There are 32 distinct symbol choices for each symbol in the data message. Therefore, each symbol represents 5 bits. The 32 symbols are represented by a single 31 bit maximum length sequence (with one additional bit added) and all possible 31 left-cyclic shifts of that same sequence.

2.2.7 Data Message - Coded

In order to improve the channel performance an outer Reed-Solomon (R-S) code is added to the data link. The R-S code, is an error control technique for use with M-ary channels. It has as large a minimum distance (d) (a measure of code power) as any other code of the same rate, and which can correct burst as well as random errors. Each codeword of the Reed-Solomon outer code is comprised of a sequence of the M-ary symbols from the inner code. The parameters of a codeword are (N,K) where K is the number of information symbols and N is the total number of symbols. One or two symbols may be computed as additional parity symbols, before this sequence is R-S encoded to improve undetected message error rate. Figure 5 illustrates the coded message format. The code parameters selected where M=32, are (N,K) = (31,15). For this code, d=17 and $2t + s \leq 16$. Any combination between 8 errors (t) and 0 erasures (s), and 0 errors and 16 erasures satisfying $2t+s \leq N-k$ will be corrected in decoding.

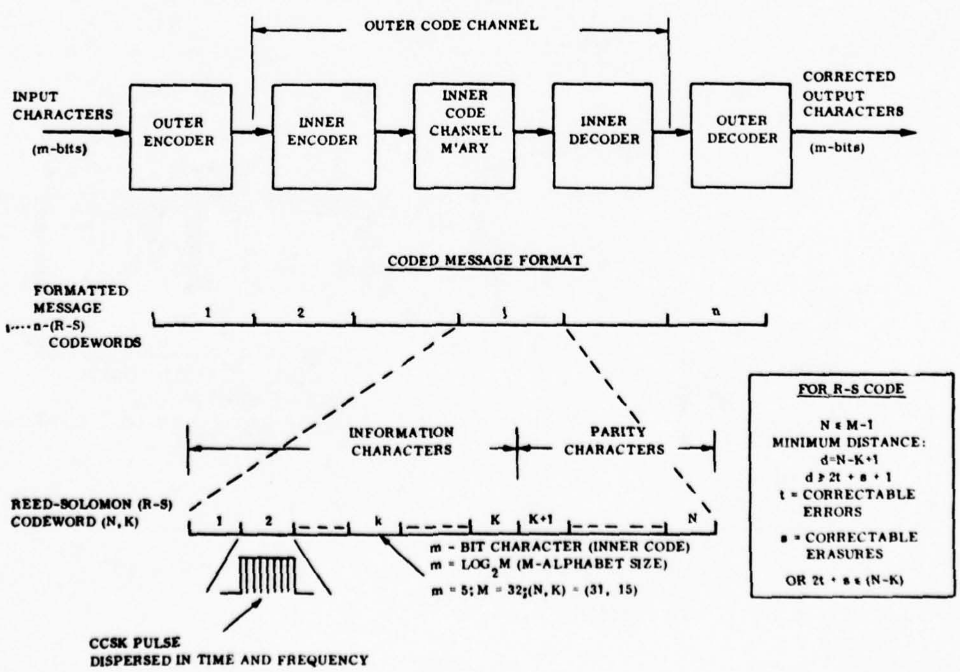


Figure 5. Concatenated Coded Channel and Coded Message Format

2.3 MODULAR CHANNEL CONCEPT

Having briefly described the fundamental elements of the basic JTIDS signal used in the DTDMA architecture, we proceed to a definition of the modular channel concept as follows:

- META-CHANNEL: The contiguous stream of all Basic Event Intervals forms by definition, a Meta-Channel. Conceptually, it may be viewed as the multiplexing of many T/F patterns (or T/F channels) each with mutually exclusive Basic Event Intervals and an associated channel event rate. The total set of channel event rates equals 78,125 Hz. Different Meta-Channels are distinguished by different contiguous (total) T/F patterns and different basic event codewords. For example, different codes of the day, or simple logical hardware changes, can create different Meta-Channels.

- BASIC CHANNEL (BC): Of fundamental importance is the particular T/F channel denoted as the Basic Channel. The 78,125 events of the Meta-Channel are subdivided into 512 Basic Channels. Therefore, each BC has an event rate of about 152.5 Hz (78,125/512). In general, these events may be divided and broadly classified as either source (Source of Communication, e.g., transmitting platform), synchronization events or message events. A typical division is normally 5 Hz and 147.5 Hz, for synchronization and message (data) respectively. For this choice, the Basic Channel is considered to be comprised of a Basic Low Event Rate Subchannel (BLS) and a Basic High Event Rate Subchannel (BHS).

- COMPOSITE CHANNEL (CC): Extending the building block approach, in which a Basic Channel is created from sets of Basic Event Intervals, a Composite Channel (CC) is created from sets of one or more Basic Channels. Combinations of more than one Basic Channel form higher event rate channels which are used as higher order building block channels to form complete function channels (e.g., Digital Voice, Command and Control, Precision Ranging, etc.). In general, composite channels support one or more services (e.g., synchronization or data) and each CC is treated as a basic entity by the JTIDS II/DTDMA Terminal.

- FUNCTION CHANNEL (FC): When a channel is designed to support a specific user function (i.e., Link 4A, Link 11, or High or Low Rate Digital Voice, etc.), it consists of sets of Basic Channels and/or Composite Channels which are organized to support the required numbers of subscribers with a prescribed level of channel capacity. The totality of these building block channels is called a Function Channel (FC). Furthermore, the component modular portions of the Function Channel are called Function/User Subchannels and have a one-to-one correspondence with the modular terminology of BC or CC. As will be shown, when illustrative function channels are described (paragraph 2.3.2), this leads to uniformity and clarity in describing and defining a single Function Channel.

Figures 6 and 7 summarize these basic channelization concepts for basic channel, function channel and multi-function channel constructions.

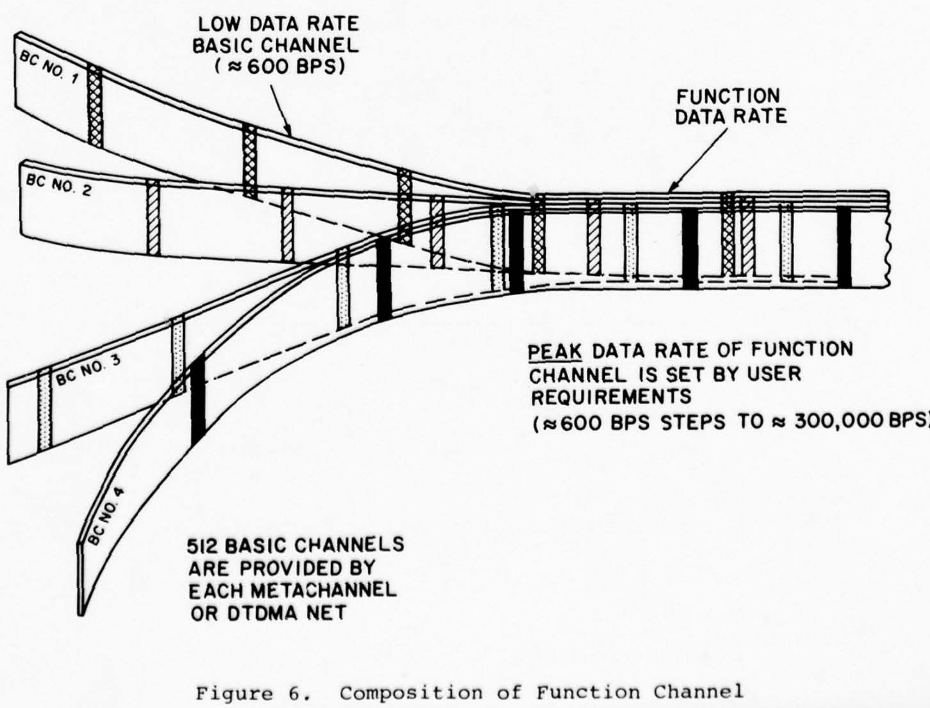


Figure 6. Composition of Function Channel

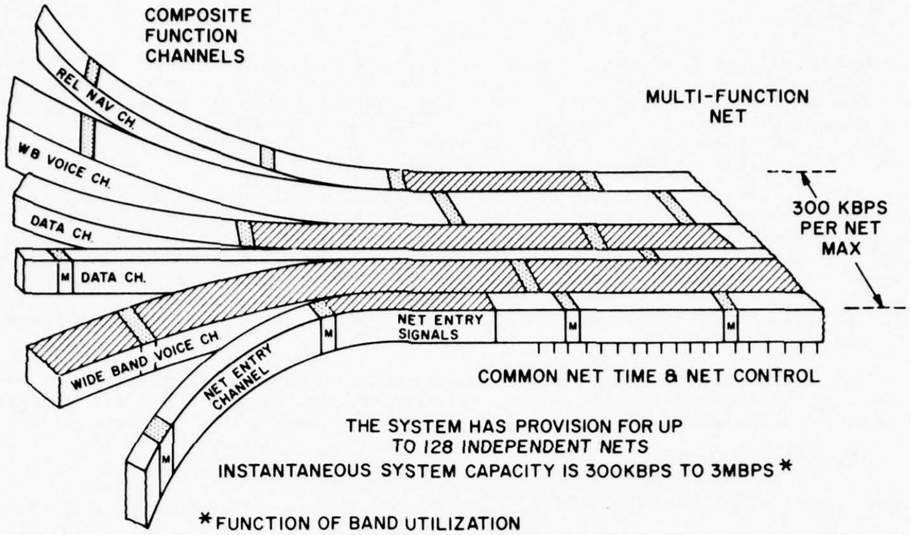


Figure 7. Composition of Multi-Function Net

From this channelization concept we note that the basic signal and channel architectures of JTIDS II DTDMA differ from the conventional TDMA structure. The basic message structure and access techniques highlight this fundamental difference.

The TDMA access technique divides time into slots with each slot containing all of the pulse signals comprising a single JTIDS I message. Short, dense bursts of pulses convey individual messages, and efficiency of time utilization favors the use of relatively long messages to maximize the ratio of message time to guard time (Figure 8 (a)). Synchronization pulses always precede the message.

DTDMA pseudo-randomly distributes the pulses conveying a message over a longer time interval, interleaved with pulses conveying other messages, without guard time (Figure 8 (b)). The low duty structure of any one message, together with the pseudo-randomly varied pulse parameters of the interleaved messages, reduces pulse interferences to a negligible level. The absence of guard time provides efficient time utilization with both short and long messages. Synchronization pulses (S) are distributed throughout the message.

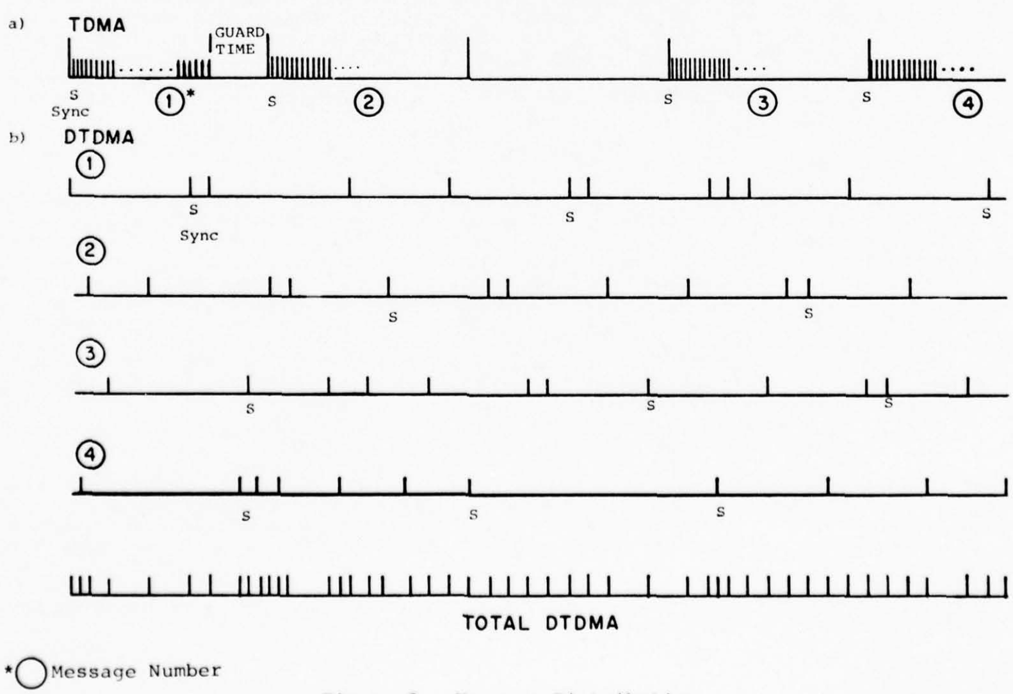


Figure 8. Message Distribution

39-2

Table 1. JTIDS II Tactical Terminal Functions and Capabilities

DTDMA FUNCTIONS

2.3.1 Generic Channel and Signal Types

Two important characteristics for further channel definition are: the types of channel access afforded to a user terminal, and the types of signals which are transmitted and received on the various subchannels of a Function Channel. Channel access introduces the concepts of: message starts (M), and open and closed communities of subscribers. Message starts (M) refer to those opportunities in time at which a message may be initiated for transmission by a given terminal (message access time).

Open and closed channel types are distinguished by the operational rules established to permit subscribers to access a Function Channel. An Open Channel is defined as a channel which is available for access at its associated M's by any subscriber without specific allocation. Thus in general, an unlimited number of subscribers can be supported by an open channel (not necessarily simultaneously). A Closed Channel is defined as a channel which is available for access only through specific subscriber allocation. Connectivity for this type of channel is limited to a finite set of subscribers, a priori known to the net community.

A useful characteristic for categorizing Function Channels is in terms of the type of access afforded to the subscribers. Four generic channel access types are pertinent to the channel design of the JTIDS II/DTDMA. They are as follows:

- Assigned Access Channel (AAC)
- Command Access Channel (CAC)
- Demand Access Channel (DAC)
- Scheduled Access Channel (SAC)

The channels may be further classified by whether they support an open or closed community of users. To indicate which of these is applicable an O or C is added to the acronym, e.g., DACO and DACC for open and closed demand access channels, respectively.

2.3.1.1 Assigned Access Channel

The Assigned Access Channel is the basic DTDMA channel where tracking of users is performed. Unique basic channel assignments are provided for each transmitting terminal. The community of users operating on that channel are all continuously tracking each other to establish expected time of arrivals for all pulses (See Section 2.3.2.1).

2.3.1.2 Command Access Channel (CAC)

The Command Access Channel (CAC), in general, is a closed type of channel and consists of a control channel and a reply channel which supports a control terminal and multiple reply, or controlled, terminals. Many options are available. The control terminal and all of the reply terminals may time-share a common CAC data subchannel, with dedicated unique CAC synchronization subchannels assigned to each terminal for supporting offset tracking of a closed community of users. Alternatively, the control and reply data subchannels may be distinct and either the reply terminals time-share the reply data subchannel or separate, assigned, reply channels may be used by each reply terminal. In all cases, the start of a message-pair (control-reply) is under the command of the control terminal and specific Message Starts (M) are associated with reply messages to minimize the control terminals search time for these replies.

2.3.1.3 Demand Access Channel (DAC)

The Demand Access Channel (DAC) may be shared by many terminals on an as-needed basis. Two types of DAC channels are distinguished by their capability to service either an unlimited, or open, community of terminals, or a limited, or closed community. For each type, procedures are provided for determining if the channel is busy, and, if so, selecting an alternate channel, or interrupting a message in progress. In an open DAC channel, message start opportunities are uniquely associated with a common synchronization subchannel which is searched by all users for message start signals or for busy signals at programmed event times. There is no need for user source synchronization tracking channel. In a closed DAC channel, message start opportunities are associated with the Basic Channel of each user, and each Basic Channel is monitored for starts. This channel requires user source synchronization tracking (See Section 2.3.2.2).

2.3.1.4 Scheduled Access Channel (SAC)

The Scheduled Access Channel is a shared channel on which messages are transmitted according to a prearranged organization. Message starts are regular and correspond to the beginning of each time slot. Transmissions are time orthogonal for this common channel because guard times (for propogation to 300 nm) are included at the end of each time slot (See Figure 8).

With these basic definitions of the generic channel types we present two illustrative channel designs for JTIDS II/DTDMA.

The functions described above have been specified for the JTIDS II Tactical Terminal. It should not be construed that the design is restricted to these functions or channels. The software programmable channel architecture of DTDMA permits a specific channel with its unique data rate, message length, or access protocol, to be readily included in the terminal. Design boundaries to be considered are the overall capacity of the terminal as limited

2.3.2 Specific Channel Implementation

We present two examples of the many that have been developed as part of the DTDMA channel catalog (software).

2.3.2.1 Assigned Access Channel - Closed (AACC)

The Assigned Access Channel is the fundamental type of user tracking distribution channel in JTIDS II/DTDMA. One of a set of Basic Channels or one of a set of Composite Channels is assigned uniquely to each terminal in a net for its transmission only and a set of access times are uniquely associated with one Basic Channel for each terminal. These access times, or Message Start Opportunities, serve the same purpose as the start of a time slot in a conventional TDMA system. However, they provide greater flexibility than the latter since there is no guard time and message may be receive simultaneously on several Composite Channels.

Figure 9 illustrates the AAC closed type of channel, where the fundamental channel is the Basic Channel. For a net of U participants, each would be assigned a Basic Channel, or a multiple number of Basic Channels, dependent on the anticipated loading for the source terminal. After the terminals have accomplished a net-entry procedure, each terminal continues to transmit user source synchronization pulses (C) which are received and tracked by all other terminals. By tracking user time offsets (opening a receiver gate on a Basic Channel over when an event is scheduled and expected), the receivers in a terminal may be time-shared between all of the Basic Channels currently in use. These coarse synchronization pulses are transmitted for source synchronization, at some low rate, whether or not a message is in progress. On this channel, there is, in general, no restriction on message duration, and both formatted or unformatted types may be used. Message start signal sequences (MST) precede data transmissions in order to keep message false alarm rates acceptable. A fine synchronization preamble (F) is used to establish synchronization accuracy sufficient for data demodulation and is also used to update user receive time offsets. Many different messages may be transmitted and received in parallel, on many composite channels, if fine-offset track is maintained on all users. The low duty factor of all messages permits a terminal's hardware to be programmed in time, frequency and pulse code and, thus process all signals in proper real time sequence and to sort the data from all active sources.

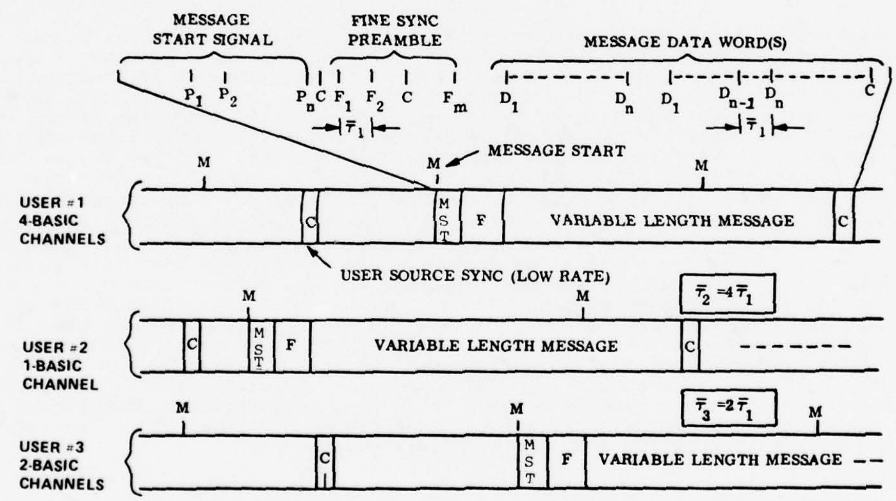


Figure 9. Assigned Access Channel (Closed)

2.3.2.2 Demand Access Channel - Open (DACO)

The Open Voice Channel (OVC) is an open Demand Access Channel (DACO) type and is designed to support high data rate digital voice, e.g., continuously variable slope delta modulation or equivalent, at a data rate of 16,000 bps. Since it is designed as an open type channel, it may be accessed by any one of an unrestricted or unlimited number of subscribers that have completed a net entry procedure (See Section 2.4) and, therefore, have become members of a (time) net that includes this channel. Additional features of this OVC are that it has an interrupt or preemption capability on the same channel or on an associated guard channel. Moreover, the same design may be used for a voice relay channel, or to support high rate digital data messages in lieu of digital voice.

Figure 10 shows that the OVC requires the union of 23 Basic Channels (BC) in order to provide a 16,000 bps data rate, including the message preamble overhead. The 23 Basic Channels from a Composite Channel which is divided into synchronization and data subchannels, and has an average event rate of 3508 Hz.

Multiple microprocessors execute the scheduling and processing of the DTDMA/TDMA JTIDS and Tacan events. The transmission or reception events for the various functions are scheduled and time ordered prior to the actual time of the event. Also completed dur-

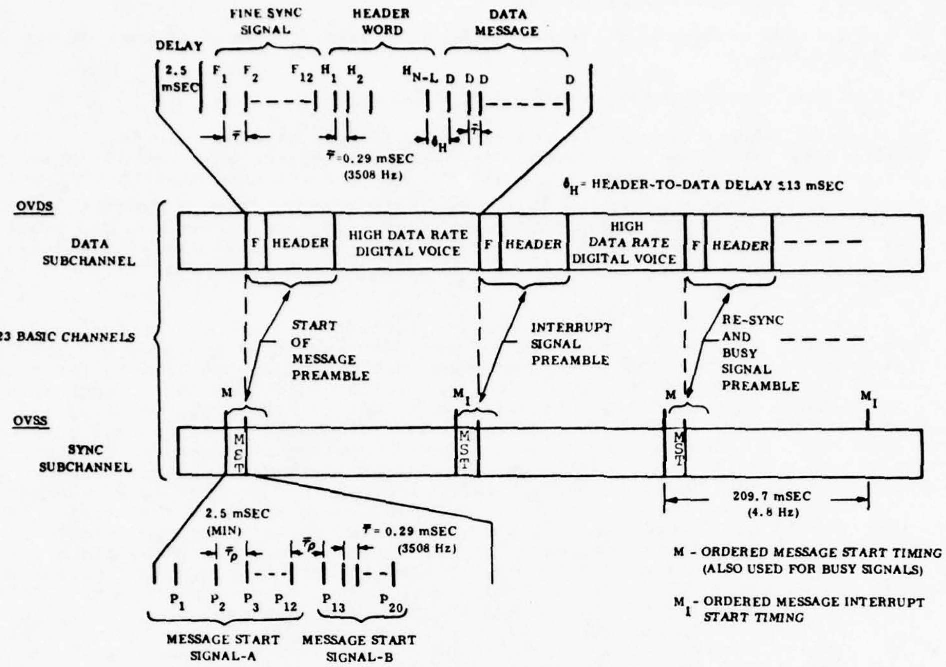


Figure 10. Demand Access Channel (Open)

An important characteristic of the OVC design is the absence of user Source synchronization signals (C). That is, it is unnecessary to track the expected time of arrivals of all potential participants in this channel. Access to this channel is on demand and only at the Message Starts (M) depicted on the OVSS, which occur every 209.7 milliseconds (4.8 Hz), and service all subscribers in the net. In order to determine if the channel is idle, each of these intervals must be voice transmission. Since user range-time is not tracked, the M signal is designed into two parts, MSA and MSB. The MSA signals correspond to a set of 12 Basic Events, selected from the sequence of channel events occurring at a rate of 3508 Hz, which satisfy a minimum time constraint of 2.5 milliseconds separation between events or pulses. The MSB signal is used to confirm the presence of a true MST signal and to decrease the false alarm probability of the MSA detection. After a 2.5-millisecond delay following the twelfth MSA pulse, the next eight Basic Events generated at a 3508 Hz rate on the 23 Basic Channel comprise the MSB signal. This narrow tracking gate is used on the MSB since the range uncertainty has been reduced by means of the MSA detections. The fine synchronization signal (F) is transmitted using the next 12 Basic Events generated on the 23 Basic Channels. Reception of this signal provides the TOA accuracy necessary to demodulate data. The last part of the message preamble and the next signal sent on the OVDS is the header word which contains the originator's ideentification number, his priority level, type of message to follow, etc. After header signal processing, the high data rate, uncoded, digital voice pulses are sent on the OVDS at an event rate of 3508 Hz which is sufficient to support 16,000 bps voice.

2.4 SYNCHRONIZATION

A comprehensive treatment of synchronization as applied in DTDMA is beyond the scope of this paper. We highlight the more salient features.

The system is designed to permit a terminal to extract navigation and communications (Data) as well as synchronization (time) information from basic signal time-of-arrival (TOA) measurements.

In order to accommodate multi-channel and multi-net operations, all transmission times and reception times are referred by "time offsets" to a common terminal time which is maintained by the terminal clock. This terminal time is the user's best estimate of universal time and is updated only in accordance with disciplines established for each terminal, based on the quality of the on-board frequency reference used by the terminal clock to maintain terminal time.

Within this mechanization approach, synchronization is the process of measuring, calculating or otherwise determining the proper "time offsets" for net transmissions and receptions.

The synchronization process consists of the four stages described in the following paragraphs.

While the emphasis has been placed on the F-14 aircraft, the same general approach can apply to a multitude of platforms. Since the JTIDS II Tactical Terminal includes the Tacan function, the existing Tacan can be eliminated. Further, an advantage of the DTDMA architecture is that digital messages of virtually any format, rate, and particular access timing can be accommodated by a DTDMA Communi...

- NET ENTRY: The net entry process permits new participants to receive special Net Entry Signals. The successful reception of these signals by an entering terminal will reduce the net time uncertainty of the new member to a value equal to the maximum one-way propagation delay (2 ms); that is, after receipt of the net entry signal the external terminal knows it is within service range but does not yet know its actual range to the signal source.

- ROUND TRIP TIMING (RTT): The second stage of timing adjustment is the RTT exchange through which net time uncertainty due to unknown propagation delays is eliminated. This round trip pulse and data exchange further reduces the timing error to 15 μs.

- COARSE SYNCHRONIZATION TRACKING: DTDMA/JTIDS II permits functions to employ source synchronization where each member of the net transmits coarse synchronization pulses (C) at an approximate rate of 5 Hz. These pulses are acquired and tracked in time of arrival by all other participants. The timing error is thus further reduced and maintained to within ±1.5 μs.

- MESSAGE SYNCHRONIZATION: To provide the high level of timing accuracy necessary for data demodulation, synchronization pulses are transmitted as part of a preamble preceding each message. These Message Start Pulses are used to obtain an overall timing accuracy, as required for data extraction and precision relative navigation.

- MULTIPLE TIME-INDEPENDENT NET PARTICIPATION: Each DTDMA function channel is capable of supporting a communications net, providing a net entry procedure, and data transfer between multiple terminals. In order to participate in multiple nets, a terminal must perform a net entry procedure for each net. This can be accomplished either serially or in parallel depending upon terminal complexity.

Since the system is designed with a one-to-one correspondence between net time and transmitted signals, multi-net operation merely requires a knowledge of the time difference between the terminal clock and net time. Once this is achieved, through the net entry procedure, the terminal then creates a net time base for each operating net. These net time bases are now available for generating the correct pseudo-random bit stream for each net, which determines the full event codes for each event interval.

3. CONCLUSIONS

Distributed TDMA is a new signaling and channel design technique for anti-jam high performance digital communication. It is being actively pursued, as an attractive solution to JTIDS Phase II, in ongoing hardware development contracts at ITT Avionics for Class I Command/Control platforms and Class 2 Tactical Terminal (See Section D1, D2). In addition, ITT Avionics is involved in studies for overall JTIDS Enhancement and Class 3 Austere/Expendable terminal applications.

We have briefly reviewed the fundamentals of DTDMA. The unique channel and signal types utilized in JTIDS II/DTDMA were highlighted with specific examples.

Table 1 summarizes the functional capabilities to be demonstrated shortly with the JTIDS II/DTDMA terminals.

Table 1. JTIDS Functional Capability

<u>Communication</u>
Anti-jam two-way digital data for direct(1) Link 4(2) interface
Anti-jam digital voice
Anti-jam TDMA digital data for direct(1) JTIDS I interoperation
Anti-jam digital data for direct(1) Link 11(3) interface
Anti-jam digital data TADIL-B(3) Interface
Precision time synchronization
Simultaneous multi-net participation
<u>Navigation</u>
Conventional Tacan (rho-theta)
Precision L-Ranging and Relative Navigation
Air/air range and bearing
Inverse mode (D/F)
<u>IFF</u>
Mark X SIF transponder capability
Mark XII Mode 4 (with GFE computer)
Full diversity operation (modular addition)
(1) Transparent interface. No changes to existing hardware or software.
(2) Command and control data link
(3) Surveillance and reporting data link

39-6

REFERENCE

The JTIDS Distributed TDMA system concept is based on the premise of maximizing the full utility of time-frequency-and-code. This multi-dimensional approach simultaneously contributes to efficient multiple access, anti-jamming, and low probability of exploitation (LPE) system solutions. The multiple access problem is solved by pseudo-random time and frequency hopping, coupled with interleaved channelization which utilizes low duty signal structures. This random access technique is particularly suited to a C^3/ICNI system as it provides intra-system and related function interference rejection.

At the same time, the combination of random access and low duty cycle permits the sharing of transmitters, receivers, and general signal processing resources for the various C-N-I functions.

The A/J problem is simultaneously solved by the hybrid DTDMA combination of the above time/frequency hopping modulation phase coded spread spectrum techniques applied to the signal pulse. There is a desired by-product of this solution to the A/J problem: The increased pulse modulation bandwidth needed for A/J provides an associated high resolution-high accuracy time-of-arrival measurement capability useful for precision ranging and navigation.

The LPI aspect of the DTDMA system is enhanced with the use of time and frequency ambiguities since much of the signal predictability is removed by the signal and channel design.

The statistical orthogonality that is designed into the DTDMA system provides not only a solution to the above C^3/ICNI problems, but also minimizes the need for more rigid organizational aspects, normally associated with conventional TDMA systems.

The promise of flexibility multi-function and multi-platform C^3/ICNI applicability, and high performance anti-jamming, secure, make DTDMA a very attractive candidate for JTIDS Phase II.

GENERAL REFERENCES

1. ITT Avionics Division, JTIDS Phase II-Advanced Development Evaluation - Task I Report, Feb. 1976, No. NADC 75296-60, Volume I (CLASSIFIED).

2. ITT Avionics Division, JTIDS Phase II-Advanced Development Evaluation - Task II Report, Nov. 1976, No. NADC 75296-60, Volume I (CLASSIFIED).

3. Cohn, C.R., Magnavox Research Laboratories, "Spread Spectrum Techniques," IEEE Press, 1976.

4. Dixon, R.C., "Spread Spectrum Techniques," IEEE Press, 1976.

5. Berlekamp, E.R., "The Development of Coding Theory," IEEE Press, 1974.

6. Peterson, W.W., "Error Correction Codes," MIT Press, Wiley, 1961.

7. Siam, J., "Reed-Solomon - Polyomial Codes Over Certain Finite Fields," Volume 8, June 1960.

ACKNOWLEDGMENTS

The basic word described herein was developed through a combination of ITT activities and Government contracts. The most recent of these, under which this work is being carried forward, is Naval Air Development Center, Contract N62269-76-C-0105. **DTDMA is not considered by the Department of Defense to be an active candidate for Phase II of JTIDS.**

SEEK BUS OPERATIONAL DEMONSTRATIONS

BY

W. F. LYNCH
The MITRE Corporation
P. O. Box 208

JTIDS II/DTDMA - COMMAND AND CONTROL TERMINALS

BY

JACK RUBIN AND JOSEPH HEINEN
ITT AVIONICS DIVISION
CNI SYSTEMS DEPARTMENT
500 WASHINGTON AVENUE
NUTLEY, N.J. 07110

SUMMARY

ITT Avionics Division is currently under contract with the JTIDS Joint Program Office/Naval Air Development Center to develop command and control terminal employing the JTIDS II/DTDMA architecture (See Section IVB). During JTIDS Phase I, TDMA was introduced to handle total connectivity digital communication and relative navigation. Distributed TDMA, an approach to JTIDS Phase II (JTIDS II), expands this capability by providing increased data rate capability and additional C-N-I functions. DTDMA also develops a channelization architecture to permit command and control terminals to participate concurrently on multiple time independent functions or networks which have the flexibility to be structured and organized in a manner to efficiently meet the broad range of command and control operations in tactical environments.

1. INTRODUCTION

The ITT Avionics JTIDS II/DTDMA Command Terminal described in this section provides specific communication, navigation and identification (CNI) functions which are selected to match operational requirements of JTIDS Phase II. The DTDMA Command and Control Terminals are intended for use by all four U.S. military services. A large variety of surface, and airborne command and control platforms are candidates for these terminals. Examples are the CVA ships, the E-2C aircraft and ground-based control centers.

For the Command and Control Terminals under development, the sizing and programming of CNI capacity results from an analysis of tri-service military command and control operations and concomitant requirements of command terminals for specific CNI services.

These terminals feature anti-jamming, secure digital voice and data, precision ranging and relative navigation, conventional Tacan navigation, and the IFF functions provided by the AIMS Mark XII transponder.

2. COMMAND AND CONTROL TERMINAL OVERVIEW

A JTIDS II/DTDMA terminal can be considered as a CNI processor which is sized and programmed to meet a specific set of user requirements. Unlike conventional single-function units where the terminal's commonly-used name (Tacan Set) conveys an understanding of its functional characteristics, a JTIDS II terminal can range in complexity from a 16-net, multi-function command terminal of greater than 150 kilobits per second (kbps) digital data capacity to a single-function, expendable missile terminal of less than 2 kbps average data capacity. (Note: These are terminal capacities. The system capacity is in excess of 300 kbps.)

Functional programs are selected from a JTIDS II/DTDMA channel catalog, which to date includes anti-jamming 16 kbps digital voice, fully compatible (with existing system) TADIL A (Link 11), TADIL B, TADIL C (Link 4), conventional Tacan/IFF, JTIDS I TDMA, and relative navigation. As new operational requirements arise, new functional channels may be configured to use the DTDMA (distributed TDMA) access technique, unique to JTIDS II because of its flexibility in efficiently accommodating a wide range of message lengths and rates.

Terminal sizing includes independent assessments of required transmitter and receiver data rate capabilities. The low duty DTDMA distributed signal structure permits concurrent transmission and reception on independent nets. A terminal's transmitter duty factor is, therefore, sized to that specific terminal's expected transmission requirements. The terminal's receiver system is independently sized to meet reception requirements, including monitoring of guard or other channels for extraction of pertinent tactical COMM-NAV-IDENT information.

The communications functions provided in the Command Terminal include secure, anti-jamming digital voice communication and secure, anti-jamming digital command and control data links. The digital voice communication is the standard 16 kbps continuous variable slope delta (CVSD) modulation for full Tri-Tac compatibility and interoperability. Either open (unlimited participation) or closed (limited participation) DTDMA channels may be selected (See Section IVB). Relay of digital voice transmissions in the same formats is also provided.

'During the Adriatic part of the E-3A

DTDMA command and control data channels handle the standard Link-4A (TADIL-C) digital data in a transparent manner for full compatibility with current NTDS/ATDS operations. Both one-way and two-way links are provided to accommodate the four Link-4A functions (Way-point insertion, CAINS update, NTDS/ATDS command and control and the automatic carrier landing system (ACLS)).

Transparent DTDMA data channels are provided for Link-11 (TADIL A), and TADIL-B digital data. Automatic two-way links accommodate the exchange of Link-11 tactical data among participating platforms.

A JTIDS I TDMA channel provides interoperability with JTIDS I terminals. Those portions of the JTIDS I Interim Joint Message Specification (IJMS) required for net entry, slot assignment/reassignment position reporting, and Round Trip Timing (RTT) are included. This permits the use of this channel for position reporting and participation in a relative navigation community on the JTIDS I TDMA channel.

The navigation functions of the Tactical Terminal include conventional Tacan (in accordance with MIL-STD-291C) and secure, anti-jamming relative navigation. The Tacan interrogator/receiver functions include full air-to-ground range and bearing service, air-to-air range and bearing service, and receive-only bearing service. Relative Navigation, which includes the near optimal blending of ranging data with on-board inertial navigation data (if the latter is available), can be performed with either a DTDMA channel or the TDMA channel, as appropriate to the particular relative navigation community.

The identification (IFF) functions provided are the conventional AIMS Mark XII transponder modes (1, 2, 3/A, C and 4), complemented by the secure, anti-jamming friend identification capability inherent in the use of JTIDS position reporting.

3. SYSTEM HARDWARE/SOFTWARE

The CNI requirements of most candidate platforms can be effectively met with three types of terminals providing three levels of participation in the JTIDS II System. The three types in order of reducing levels of participation are the Command and Control, Tactical and Expendable Terminals.

The Command and Control Terminal under current development has a maximum data rate capability of 150 kilobits per second. This capacity can be divided among the following communications systems: Link-4A, Link-11, JTIDS I (IJMS), TADIL B, Precision Ranging and Rel Nav, Open Voice, and Closed Voice. The IFF transponder function and Tacan navigation are also provided. While participating in up to sixteen time independent communication nets concurrently at throughput rates up to the maximum capacity, the terminal obtains the highest anti-jamming performance afforded by the net organization and signal structure of the DTDMA system.

A simplified block diagram of the Command Terminal is provided in Figure 1 which shows the five major blocks of the terminal, Power Amplifier (PA), the Receiver/Transmitter (R/T), Signal Processor (SP), Terminal Processor (TP) and the Control and Display Units.

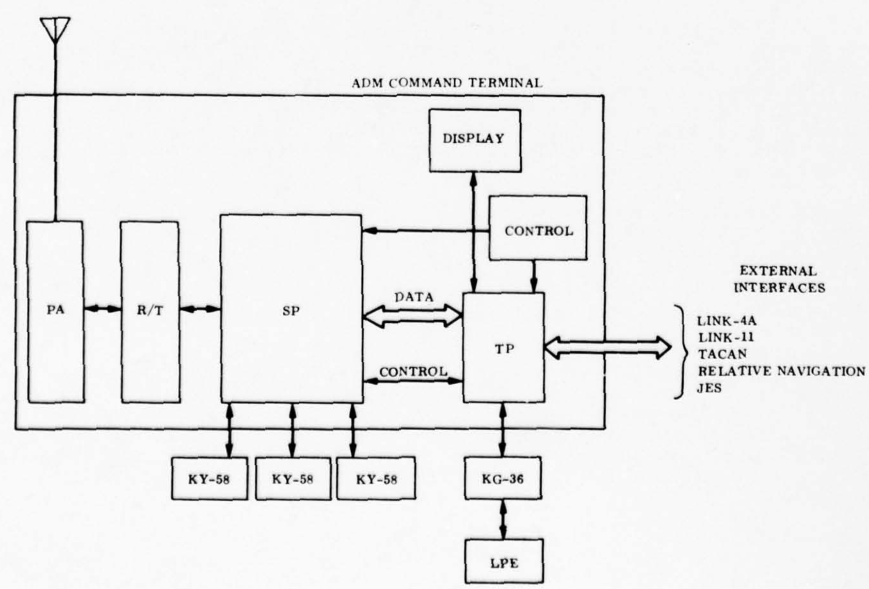


Figure 1. JTIDS II/DTDMA Command and Control Terminal, Block Diagram

Table I is a summary of pertinent technical characteristics of the Interim SEEK BUS terminal.

TABLE I

3.1 POWER AMPLIFIER UNIT

The Power Amplifier Unit (PA) provides the power gain and modulation control for all transmitted pulses. The basic unit is a broadband (255 MHz) solid-state amplifier which takes the 1 to 2 mW input signal from the R/T unit and supplies over 1 kW output across the entire band. It contains a single RF driver module to drive the two Power Amplifier modules. The basic PA module consists of six broadband power transistors with 800 watt peak power output.

Bandpass filters are used in both the transmit and receive paths and provides controlled transmit spectrum performance and enhanced spurious response rejection.

The unit contains an amplitude modulator which provides the basic control for power level and various JTIDS modes of JTIDS I and II, Tacan Gaussian and IFF rectangular pulse shapes.

3.2 RECEIVER/TRANSMITTER UNIT

The basic unit is a multi-function, dual-conversion heterodyne receiver, capable of simultaneous operation on four independent channels. The spread-spectrum continuous phase shift modulated (CPSM) JTIDS pulses, as well as conventional Tacan Gaussian and rectangular IFF, are processed.

The basic transmission service for the CPSM spread spectrum JTIDS pulses is provided by the CPSM Modulator and Up-Converter modules. The frequency inputs are provided from the frequency source and frequency synthesizer modules.

There are four high speed synthesizers in the R/T Unit. Phase locked to the systems reference standard, these synthesizers are capable of switching in 1-MHz steps across the full 255-MHz band.

A Quadrature L-Band to 70 MHz converter uses the synthesizer outputs with the 280-MHz output of the frequency source module to provide dual down-conversion of the incoming L-Band signal to the final IF frequency of 70 MHz.

The four IF amplifiers (one per channel) provide 80 dB of gain and each has both hard limited and log video outputs. IFF reception is accomplished by using the log video output of one channel. For Tacan reception, a combination of the hard limited and log video output of one channel is used. Matched filter detection is used to correlate the spread spectrum pulses. A fully programmable Surface Acoustic Wave (SAW) device is fed by each of the IF amplifiers. An erasure threshold, modified maximum-likelihood scheme is used to detect the correlator outputs.

3.3 SIGNAL PROCESSOR UNIT

The Signal Processor is the basic controller for the entire system. Before describing the functions of the various modules of the Signal Processor, a brief discussion of the data flow is presented.

Figure 2 is a data flow diagram for the transmit and receive modes. The basic data enters the flow at point (1). At point (2) the data exits the data storage area. To achieve high levels of anti-jamming performance, several additional levels of protection are added. The Reed-Solomon Encoder accepts the data at point (3), provides Reed-Solomon encoding as required, and sends the coded data to the modulator at point (4), where each 5-bit character is impressed on a 32-chip digital word. The data modulator then sends the data modulated-coded data back to the PR source at point (5). At this point, the PR source performs two functions. The first is to provide a pseudo-random (PR) bit stream to the frequency-time hopping control at point (7). This PR bit stream provides a pseudo-random carrier frequency and time hopped transmission schedule at point (8). Second, the PR source processes the encoded data received at point (5) by a PR mixing process wherein the data bits are added modulo-two to an equal number of PR bits. This process is designated as PR spreading. The PR spread data is sent to the RF modulator in the transmitter at point (6) and phase modulates the RF carrier at the time of transmission.

The end result of this process is RF pulses at L-Band at point (9). These pulses have high anti-jamming properties and can be characterized as follows:

Data Collection/Buffering	Introduced between Points (1) and (2)
Error Protected:	Introduced between Points (3) and (4)
Data Modulated:	Introduced between Points (4) and (5)
PR Spread:	Introduced between Points (5) and (6)
Time Frequency Hopped:	Introduced between Points (6) and (9)

The radiated pulses from point (9) are remotely received at point (10). A PR bit stream at point (11) provides instructions to the time/frequency dehopping control to correctly de-hop the received pulses at point (12). The resulting output at point (13) is provided to the data detector which first strips (removes) the PR spreading code at

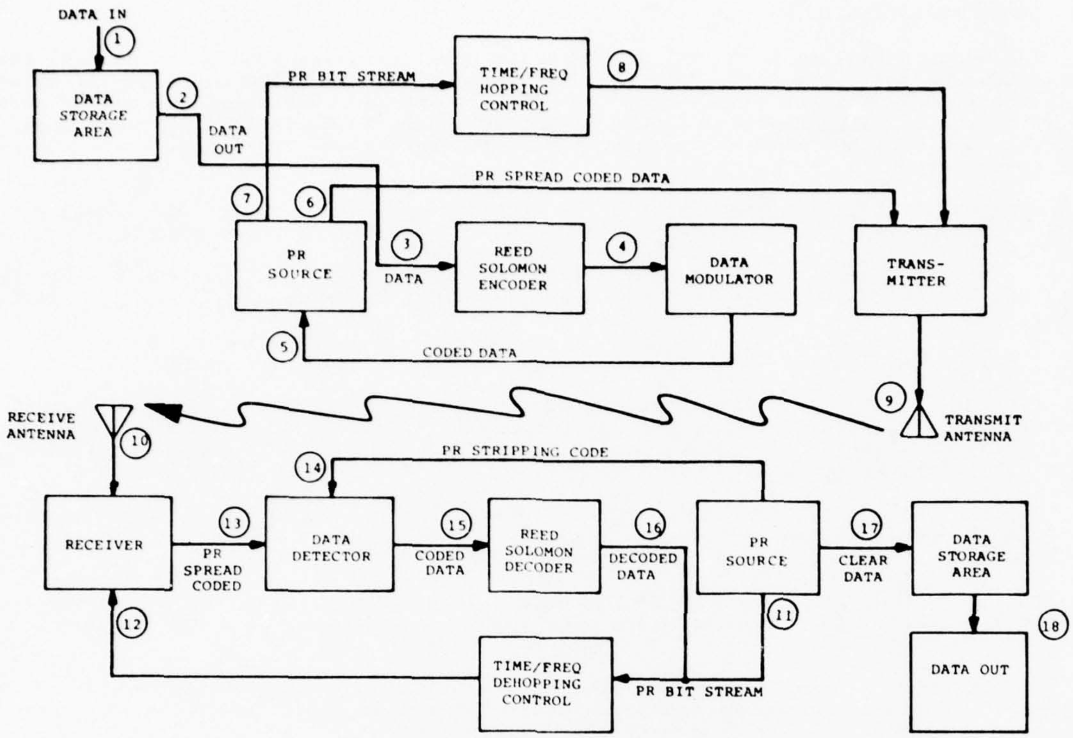


Figure 2. Transmission and Reception Data Flow

point (14), and then demodulates the 32 chip data word. The resultant 5-bit data character then enters the Reed-Solomon decoder at point (15), and the decoded data resulting at point (16) enters the data storage area at point (17) and is buffer stored for dissemination to the required interface at the appropriate time at point (18).

The hardware elements of the Signal Processor are organized into major functional groups as shown in Figure 3. A brief description of all major SP modules and their function is provided in the following paragraphs.

The unit can be broken down into major functional modules. We highlight the major groupings below.

3.3.1 Time Scheduling

This group includes the Schedule Processor, Time Ordered Lists and Terminal Clock and Composite Storage Modules (Figure 3). For a Distributed TDMA architecture there exists

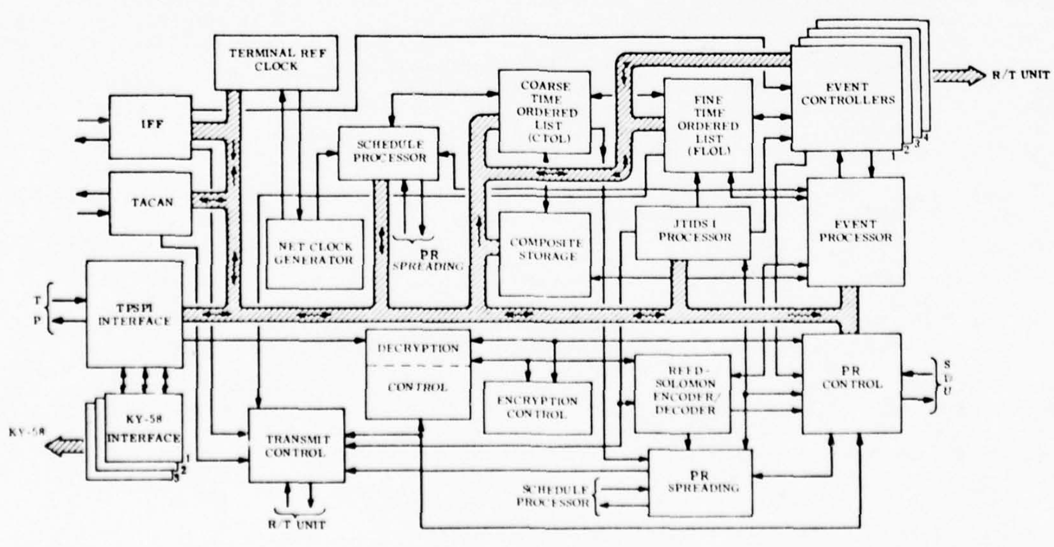


Figure 3. Signal Processor, Block Diagram

a pseudo-random time between the basic transmission and reception events. Within the Signal Processor unit, the basic function of the Schedule Processor is the ordering of all transmission and reception events to ensure that the required event sequence is monitored. The Schedule Processor generates these time ordered events and loads them into a Coarse Time Ordered List. The Coarse Time Ordered List is the mechanism whereby events created by the Schedule Processor are placed in the correct time order. This list formats time parameters of scheduled events and stores them until near real time before passing them to the Fine Time Ordered List. The Fine Time Ordered List assembles specified event parameters, determines the priorities of the events, and assigns these events to one of the four receiving channels or to the transmitter as required.

3.3.2 Pseudo-Random Control

This grouping includes the PR Spreading, Encryption/Decryption and PR Control Modules (Figure 3). This section contains the circuitry to control the pseudo-random bit streams to service both TDMA (JTIDS I) and DTDMA (JTIDS II) for the functions.

This module controls the pseudo-random phase and frequency elements of each JTIDS pulse. On transmission, the module provides the basic phase code for the transmitted pulse and the associated carrier frequency. On reception, the module provides the reference PR code to be used for reception and the associated carrier frequency.

Additional control circuitry is responsible for controlling the flow of data through the decryption/encryption process. The Decryption Control compiles the encrypted data and outputs it to the Terminal Processor for final distribution.

The Encryption Control extracts the data from the Terminal Processor and deposits it in the Reed-Solomon Encoder/Decoder section.

3.3.3 Transmission/Reception Control

This grouping includes the Event Controllers, Event Processor and Reed-Solomon Encoder/Decoder Modules (Figure 3). For transmit events, the Transmit Controller Module receives message from the data memories. These transmit events are screened to adjudicate any conflicts before passing the data to R/T Unit for actual transmission.

For received events the Event Controller takes the time ordered schedule and assigns one of four available receivers. Appropriate phase code and frequency code are supplied to one of the four RF channels to complete the reception process. The received time of arrival of the pulse (TOA) is returned from the R/T Unit to the Event Controller.

The Event Processor receives data and control information from the Event Controller. The Event Processor adjusts the appropriate tracking loop on the basis of the received time error. The Event Processor also determines if the received message has been Reed-Solomon encoded. Non-encoded messages are passed directly to the Decryption Control circuit. Encoded messages are passes via the Reed-Solomon Decoder, to the Decryption Control.

Other modules provide the JTIDS I TDMA processing function, the Tacan and IFF function, Reed-Solomon encoding/decoding and basic clocking and sequence control for the unit.

3.4 TERMINAL PROCESSOR UNIT

The Terminal Processor (TP) (Figure 1) provides a centralized, flexible digital control and processing center for the ADM Command Terminal. The following major tasks are performed by this unit:

- Mode Control and Channel Configuration
- Buffering and Two-Way Data Flow Control with External Data (for example Link-4A, TADIL B, LINK-11).
- Control and Interface with Display Unit.
- Initialization and Monitoring of Built-in-Test.

The Terminal Processor contains two major elements. The first is the Computer Interface Block which contains the CX2-475 computer and associated memory in addition to the SP I/O, Display Unit I/O, KG-36 I/O and external serial interference.

The second major element of the Terminal Processor is the Platform Interface Block which provides the platform dependent interface with the NTDS and ATDS computers and associated peripherals for the Link-4A, TADIL B and Link-11 services. Additional interfaces with the Tacan control panel and analog instrumentation are also provided by the Platform Interface Block. In general, this block varies with the specific platform, whereas the Computer Interface Block remains constant.

3.5 CONTROL AND DISPLAY UNIT

The Control and Display Unit (CDU) (Figure 1) provides the direct man-machine interface in the form of an alphanumeric display, a cued keyboard for data entry and several control selector switches. Its electrical interface is with the Terminal Processor. The operator provides initialization data and mode control commands to the Terminal Processor via the Control and Display Unit and is cued in this task through the alphanumeric displays and cueing lights. The alphanumeric display is also used to display status information, Tacan data, and malfunction isolation information.

4. INSTALLATION AND INTEGRATION OF JTIDS II/DTDMA COMMAND AND CONTROL TERMINALS

The Command and Control Terminals under development have been designed to interface with typical command and control platforms. The large surface platforms are typified by the CVA installations, while the airborne command posts by the E-2C installations. Figure 4 shows a typical interconnection for the CVA. We note that, because of the distributed architecture and the channel, message and timing flexibility it offers, the interfaces are direct and transparent to the existing systems such as Link-4A (TADIL C) and Link-11 (TADIL A). No modifications are forced, and the advantages of enhanced anti-jamming, secure high performance data links are provided. Three simultaneous high rate (16 kbps) voice and one low data rate coded data/voice interface are provided.

5. PRODUCTION JTIDS II/DTDMA COMMAND AND CONTROL TERMINALS

Figure 5 shows the Command and Control Terminals as envisioned in a production configuration. The CNI Transceiver unit will be functionally similar to the existing Power Amplifier and Receiver/Transmitter unit of the current development. The basic improvements are in packaging and utilization of latest technology. The CNI Processor Unit combines the functions of the current Signal and Terminal Processor. As before, with the rapid advance of digital technology this unit represents a reasonable extension of today's technology.

6. CONCLUSION

The JTIDS II/DTDMA Command and Control Terminals will provide JTIDS Phase II requirements. Multi-channel secure, anti-jamming digital data and voice communication, precision ranging and relative grid navigation interoperability with JTIDS I TDMA as well as conventional Tacan and IFF services are offered.

The DTDMA architecture assures minimum impact transparent interfaces for providing high rate 16 kbps/low rate (2.4 kbps) digital/data voice, Link-4A (TADIL C), Link 11 (TADIL A) and TADIL B.

The promise of flexibility, multi-function and multi-platform applicability, and high performance anti-jamming, secure communications makes JTIDS/DTDMA a very attractive candidate for JTIDS Phase II. It is currently under further development.

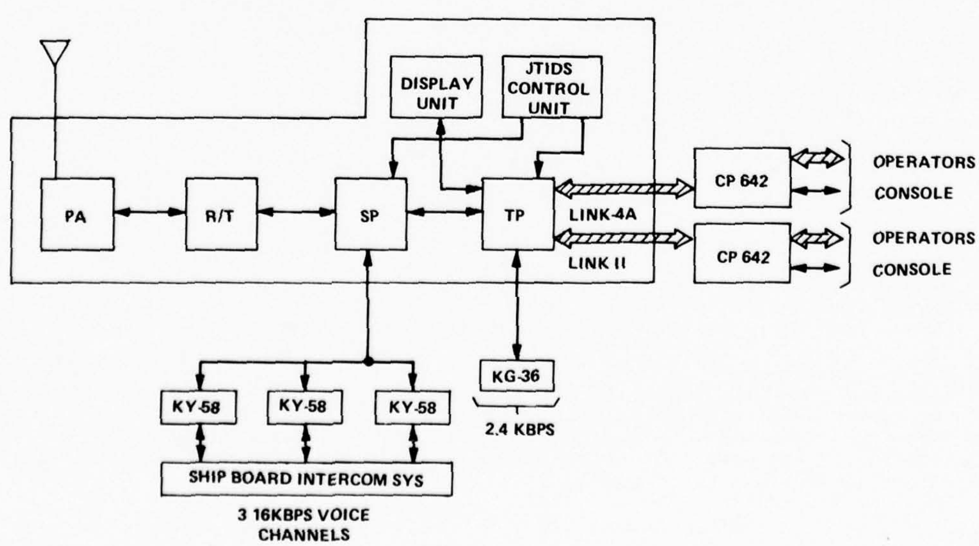


Figure 4. JTIDS II/DTDMA Command and Control Terminal - CVA Interface

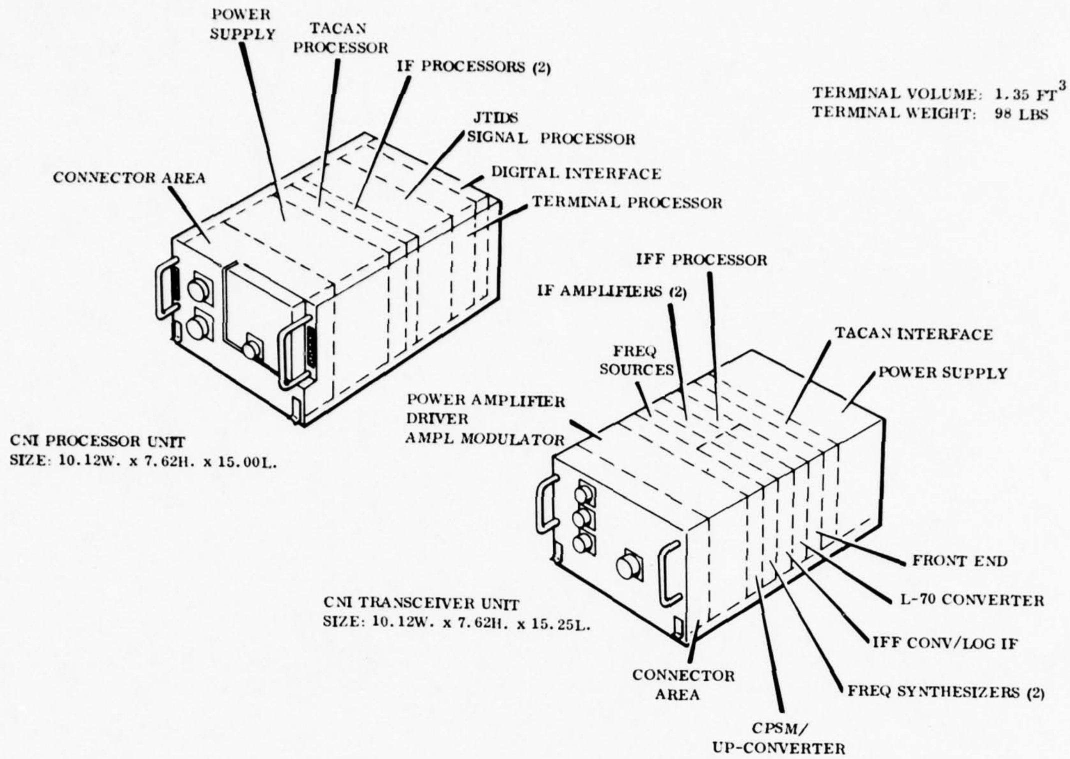


Figure 5. Production JTIDS II Command and Control Terminal (TFH Logic)

REFERENCE

Rubin, J., 1978 NATO Agardograph, Part III - JTIDS, "Distributed TDMA - An Approach to JTIDS Phase II," Paper IVB.

ACKNOWLEDGMENTS

The basic work described herein was developed through a combination of ITT activities and Government contracts. The most recent of these, under which this work is being carried forward, is JTIDS Joint Program Office/Naval Air Development Center, Contract No. N62269-76-C-0105. DTDMA is not considered by the Department of Defense to be an active candidate for Phase II of JTIDS.

JTIDS II/DTDMA TACTICAL TERMINAL

BY

JOSEPH HEINEN AND JACK RUBIN
ITT AVIONICS DIVISION
CNI SYSTEMS DEPARTMENT
500 WASHINGTON AVENUE
NUTLEY, N.J. 07110

SUMMARY

ITT Avionics Division is currently under contract with the JTIDS Joint Program Office/Naval Air Development Center to develop tactical terminals employing the JTIDS II/DTDMA architecture (See Section IVB). During JTIDS Phase I, TDMA was introduced to handle total connectivity digital communication and relative navigation. Distributed TDMA, an approach to JTIDS Phase II (JTIDS II), expands this capability by providing increased data rate capability and additional C-N·I functions.

The JTIDS II Tactical Terminal provides secure, anti-jam, digital data/voice communications and relative navigation in the DTDMA and TDMA architectures. The standard airborne Tacan and IFF transponder functions are integrated with JTIDS into the single terminal.

1. INTRODUCTION

ITT Avionics is currently under contract to the U.S. Navy and the JTIDS Joint Program Office to build JTIDS Phase II Tactical Terminals(1). These terminals accomplish multiple digital voice/data functions and relative navigation within the JTIDS II Distributed Time Division Multiple Access (DTDMA) architecture, interoperate with JTIDS Phase I terminals employing the TDMA architecture, and also perform the standard airborne Tacan and IFF transponder functions. The Tactical Terminals are to be used on small platforms including tactical aircraft and small ships for the primary purposes of providing secure, anti-jam digital communications and precision navigation. The JTIDS II Tactical Terminal functions and capabilities, its overall design, its integration and installation into airborne platforms and the projected production configuration of the terminal are described in this paper.

2. TERMINAL FUNCTIONS AND CAPABILITIES

The JTIDS II Tactical Terminal functions which were selected for inclusion in the terminals resulted from operational scenario analyses performed by the U.S. Armed Forces and the JTIDS Joint Program Office. The functions represent a consensus of which capabilities are required in a JTIDS Tactical Terminal for a tactical fighter aircraft. Although not specifically limited to a single aircraft type, the U.S. Navy's F-14 was used as the candidate for defining representative operational requirements. The Tactical Terminal is also applicable to other platforms such as the F-4, F-15, F-16, F-18 and A-7.

The terminal functions and the resultant digital data rates are summarized in Table 1. The JTIDS functions are segregated between the Distributed TDMA (DTDMA) and the TDMA architectures. The terminal yields the full anti-jam protection of the JTIDS architectures, including the spread spectrum pseudo-random (PR) phase coding, the PR fast frequency hopping common to TDMA and DTDMA and the transmission of each JTIDS pulse at random times as uniquely provided by the DTDMA architecture (See Reference).

The random transmission of pulses permits concurrent transmission and reception of messages, which in turn permits a DTDMA terminal to participate simultaneously in asynchronous (time independent) nets. The Tactical Terminal is designed to synchronize with and communicate and/or navigate on any four of 128 synchronous or asynchronous time nets. It will also simultaneously interoperate with the TDMA net.

The digital voice channels are compatible with standard 16 kbps CVSD modulated voice. Each voice channel has an interface with a KY-58 voice encoder/decoder. One voice channel can be automatically reconfigured to provide relay service to terminals beyond line-of-sight of the transmitting terminal.

---

(1) JTIDS Joint Program Office/Naval Air Development Center Contract No. N62269-76-C-0105

TOTAL DTDMA

*◯ Message Number

Figure 8. Message Distribution

Table 1. JTIDS II Tactical Terminal Functions and Capabilities

DTDMA FUNCTIONS

- One 16 kbps Direct Voice Channel
- One 16 kbps Direct or Relay Voice Channel
- Link 4A (TADIL C) Digital Data Link
  - One-Way (CAINS, Waypoint)
  - Two-Way (Command and Control, ACLS)
- Link 11 (TADIL A) Digital Data Link Relay
- Relative Navigation
- Peak Data Rate Loading Of Above Functions
  - Receive - ≈50 kbps
  - Transmit - ≈40 kbps
- Data Rate Capacity (DTDMA) Of Terminal
  - Receive - ≈70 kbps
  - Transmit - ≈50 kbps

TDMA FUNCTIONS

- Relative Navigation Interoperability
- Net Management Processing
- Round Trip Time Measurement
- Data Rate Capacity (TDMA) Of Terminal
  - Receive - ≈70 kbps
  - or
  - Transmit - ≈70 kbps

AIRBORNE TACAN (SIMILAR TO AN/ARN-84)

- Receive, Transmit/Receive, and Air-to-Air Modes
- Tacan Inverse Mode

IFF TRANSPONDER (SIMILAR TO AN/APX-72)

- AIMS Mark XII Modes 1, 2, 3/A, C and 4

One-way and two-way Link-4A digital data channels are provided for all operational Link-4A functions on the F-14. These include: (1) the Carrier Aircraft Inertial Navigation System (CAINS) on-deck alignment function and (2), the waypoint insertion function, both of which are one-way assigned access transmissions to the aircraft terminal, (3), the Naval Tactical Data System/Airborne Tactical Data System (NTDS/ATDS) command and control function and (4), the Aircraft Carrier Landing System (ACLS) function, both of which are two-way command access digital data transmissions.

The Link-4A channels of the Tactical Terminal provide transparent link operation; that is, switching of Link-4A operation from the current UHF data links to the JTIDS-II secure, anti-jamming channels does not affect the baseband processing and display functions of Link-4A operation. This transparency capability arises from the compatibility of the DTDMA Link-4A timing and data formats with the standard timing and data formats of current shipboard and airborne Link-4A computers. The DTDMA Link-4A channels use Reed-Solomon error detection and correction coding to provide an increased level of jamming immunity.

The DTDMA Link-11 relay channel also employs Reed-Solomon error detection and correction. The Relay Link-11 channel provides relay service to users beyond line-of-sight of the terminal initiating a Link-11 report.

The Tactical Terminal provides an interoperable JTIDS I TDMA or a DTDMA Relative Navigation Channel. Precision measurement of ranges between participating aircraft is used, together with available navigation data from other aircraft and from on-board inertial sensors, to achieve an accurate, three-dimensional navigation grid for multiple terminal operations.

The TDMA (JTIDS I) channel supports the full 128 time slots per second capacity of a JTIDS I net. Accesses to these time slots are scheduled. Maximum receive data rate for this channel is 70 kbps. The TDMA channel usage will typically be used for interoperability with terminals having only a JTIDS I TDMA capability. A possible scenario is participation in a JTIDS I relative navigation net, using the position reports, round-trip timing interrogations and replies, and net entry and slot assignment messages of the JTIDS I Interim Joint Message Standards (IJMS).

The Tacan interrogator function includes the standard Receive (REC), Transmit/Receive (T/R), and Air-to-Air (A/A) modes operational on tactical aircraft. In addition, the Tacan Inverse mode is provided. On aircraft equipped with a rotating type Tacan antenna, placing Tacan in the inverse mode causes the antenna pattern to rotate at a 15 Hz rate. This permits the Tactical Terminal to measure bearing to another platform transmitting normal Tacan signals.

The IFF transponder function of the Tactical Terminal is similar to that provided by the AN/APX-72 IFF set. Included are the AIMS Mark XII transponder modes (1, 2, 3/A, C and 4).

(See Figure 8).

With these basic definitions of the generic channel types we present two illustrative channel designs for JTIDS II/DTDMA.

The functions described above have been specified for the JTIDS II Tactical Terminal. It should not be construed that the design is restricted to these functions or channels. The software programmable channel architecture of DTDMA permits a specific channel with its unique data rate, message length, or access protocol, to be readily included in the terminal. Design boundaries to be considered are the overall capacity of the terminal as limited by the number of receivers, the specified transmitter duty cycle, and the real-time processing capability of the digital hardware.

3. TERMINAL DESIGN

The design approach taken with the JTIDS II Tactical Terminal was to achieve its multiple communication, navigation and identification (CNI) functions through maximum time sharing of hardware, thereby reducing the total hardware, and to maximize the use of microprocessors in lieu of discrete digital logic. The Tactical Terminal electrical design is generally similar to the JTIDS II Command Terminal with some important differences.

A major advantage of DTDMA is that with reduced data rate requirements, the number of receiver channels can be reduced while still maintaining full anti-jamming protection. Thus, the Tactical Terminal needs only two receiver channels for JTIDS whereas the higher data rate Command Terminal has four receivers. Since each receiver channel includes not only an IF amplifier but also a frequency synthesizer, a spread spectrum pulse compressor/detector, and real-time digital controlling circuitry, the elimination of receivers without loss of anti-jamming protection is a significant cost and size savings. DTDMA further permits one to configure a single receiver terminal at lower data rates, while again maintaining full anti-jam protection. This single receiver approach is feasible for JTIDS Class 3 terminals, used as Army manpacks or for missile guidance.

The simplified block diagram (Figure 1) shows that the Tactical Terminal has two time-shared receivers for JTIDS and Tacan and a dedicated IFF receiver, all serviced by a common RF front end. The spread spectrum pulses from each JTIDS receiver feed an IF pulse compressor/correlator which uses a programmable SAW pulse compressor. The SAW compressor has near optimum processing gain and therefore produces better anti-jam performance than suboptimum implementations such as digital or CCD pulse correlators. The transmitter's solid-state exciter and power amplifier outputs JTIDS, Tacan and IFF pulses each of which is appropriately shaped by the amplitude modulator. The transmitter's frequency bandwidth covers all JTIDS/Tacan/IFF frequencies. The output pulses are interleaved in real-time with a peak power of 200/800 watts for JTIDS and 800 watts for Tacan and IFF.

The terminal further contains analog and digital processing hardware for Tacan and IFF. The Tacan processing algorithms have been designed to give specified range and bearing accuracy while only utilizing a single receiver for approximately 1 percent of real time. Having highest priority, when an IFF interrogation is detected, the transmitter is instantly made available for the IFF response.

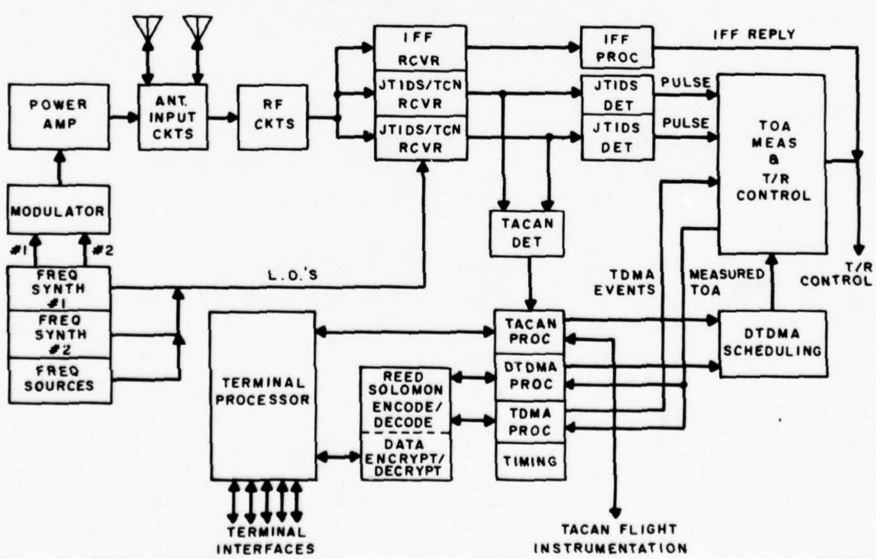


Figure 1. JTIDS II Tactical Terminal (Simplified Block Diagram)

Figure 10 shows that the OVC requires the union of 23 Basic Channels (BC) in order to provide a 16,000 bps data rate, including the message preamble overhead. The 23 Basic Channels from a Composite Channel which is divided into synchronization and data subchannels, and has an average event rate of 3508 Hz.

Multiple microprocessors execute the scheduling and processing of the DTDMA/TDMA JTIDS and Tacan events. The transmission or reception events for the various functions are scheduled and time ordered prior to the actual time of the event. Also completed during this time is the encryption and Reed-Solomon encoding of data for transmission plus processing of pseudo-random code for phase coding and determination of the RF frequency. All functions timeshare a single Reed-Solomon decoder.

The terminal processor is a general-purpose computer which executes such algorithms as relative navigation, Tacan tracking, control/display, terminal initialization, message processing and interface control.

4. INSTALLATION AND INTEGRATION OF THE JTIDS II TACTICAL TERMINAL INTO AIRBORNE PLATFORMS

The JTIDS II Advanced Development Model Tactical Terminals currently being built under U.S. Navy contract consist of three units per terminal. (The production terminal, described later in this paper, will have two units per terminal.) As shown in Figure 2, the three units of the current terminal are the:

- CNI Transceiver Unit
- CNI Processor Unit
- Interface Unit

In general, the partitioning of functional hardware conforms to the names of the units. The CNI Transceiver Unit contains all the receiving and transmitting hardware plus the IFF processor. The CNI Processor Unit is primarily the digital processing unit, but also contains the two IF pulse compressors/correlators, analog Tacan processing hardware, and a power supply. The Interface Unit contains the digital interfaces to/from other equipments, the analog signals for Tacan instruments, and a second power supply.

The Interface Unit is a form and fit replacement for the AN/ARN-84 Tacan Receiver-Transmitter and associated Signal Data Converter, which is installed on the F-14 aircraft. This approach simplifies F-14 (or any platform containing Tacan) modification since the Tacan Control, Attitude Heading and Reference System (AHRS), and Tacan instruments interconnections can be made without aircraft rewiring. Of importance is that replacement of the Tacan equipment reduces the perennial space problem associated with installing new equipments during aircraft modification programs.

The Link-4A digital interface to the ASW-27B and AWG-9 equipments in the F-14 allows the JTIDS terminal to supplant functions normally done by the existing UHF Link-4A equipment without any change to onboard hardware, computer software, or Link-4A system timing. This is referred to as a "transparent" interface. The digital Link-4A signals can be transmitted and received via JTIDS (with added encryption and anti-jam capability), instead of via the UHF Link-4A data link, without any change to the user equipments. The JTIDS equipment could also potentially eliminate the ASW-27B UHF Link-4A equipment if a relatively small amount of interface hardware is added to JTIDS.

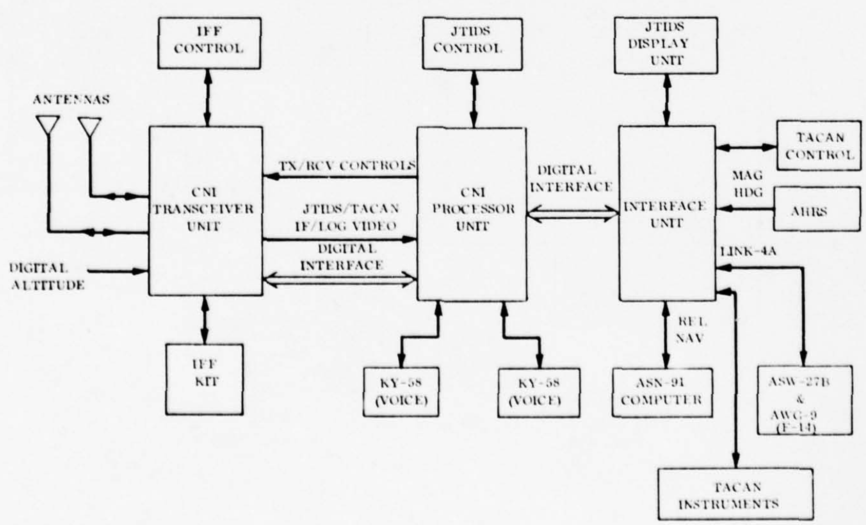


Figure 2. ADM JTIDS II Tactical Terminal Interfaces

lating or otherwise determining the proper "time offsets" for net transmissions and receptions.

The synchronization process consists of the four stages described in the following paragraphs.

While the emphasis has been placed on the F-14 aircraft, the same general approach can apply to a multitude of platforms. Since the JTIDS II Tactical Terminal includes the Tacan function, the existing Tacan can be eliminated. Further, an advantage of the DTDMA architecture is that digital messages of virtually any format, rate, and particular access timing can be accommodated by a DTDMA Communications Channel. These "transparent" interfaces with JTIDS reduce the need for changes to equipments with which JTIDS will interface. The Tactical Terminal also contains the IFF transponder function, thereby allowing the existing IFF transponder to be removed. The potential also exists for the partial elimination of multiple UHF or VHF voice sets. Finally, the integrated JTIDS/IFF and Tacan functions use the same antenna subsystem and, on most platforms, the Tactical Terminal should utilize the existing Tacan and/or IFF antennas.

5. PRODUCTION JTIDS II TACTICAL TERMINAL

The current equipment has three units per terminal. The production JTIDS II Tactical Terminal will be packaged within two units, the CNI Transceiver and CNI Processor (Figure 3). The CNI Transceiver functionally will be the same except it will contain the Tacan interface and will be implemented with the latest technology. The CNI Processor will contain not only the processing functions, but also the interface hardware and power supply which had been packaged separately in the Interface Unit.

The CNI Transceiver Unit would replace the AN/ARN-84 Tacan and provide the installation advantages described in the prior section. Both units have the same ARINC one ATR cross section (maximum dimensions of 10-1/8 inches wide by 7-5/8 inches high) and have a combined volume of 1.35 cubic feet.

6. CONCLUSIONS

The JTIDS Phase II Tactical Terminal will provide secure, anti-jam communication, coupled with conventional Tacan, precision relative navigation and IFF transponder functions.

The terminal includes two 16 kbps digital voice, voice relay, Link 11 (TADIL A) relay, and one-way/two-way Link-4A (TADIL C) channels in the JTIDS DTDMA architecture. Interoperability with JTIDS TDMA terminals is provided. Precision position location is available via the relative navigation function.

Installation and integration of the terminal is enhanced by the potential replacement of existing Tacan, IFF and UHF voice sets and the utilization of existing antennas. Transparent interfaces minimize or eliminate modification of interfacing equipment.

The production JTIDS II Tactical Terminal will be two units, the CNI Transceiver and CNI Processor, with a total terminal volume of 1.35 cubic feet.

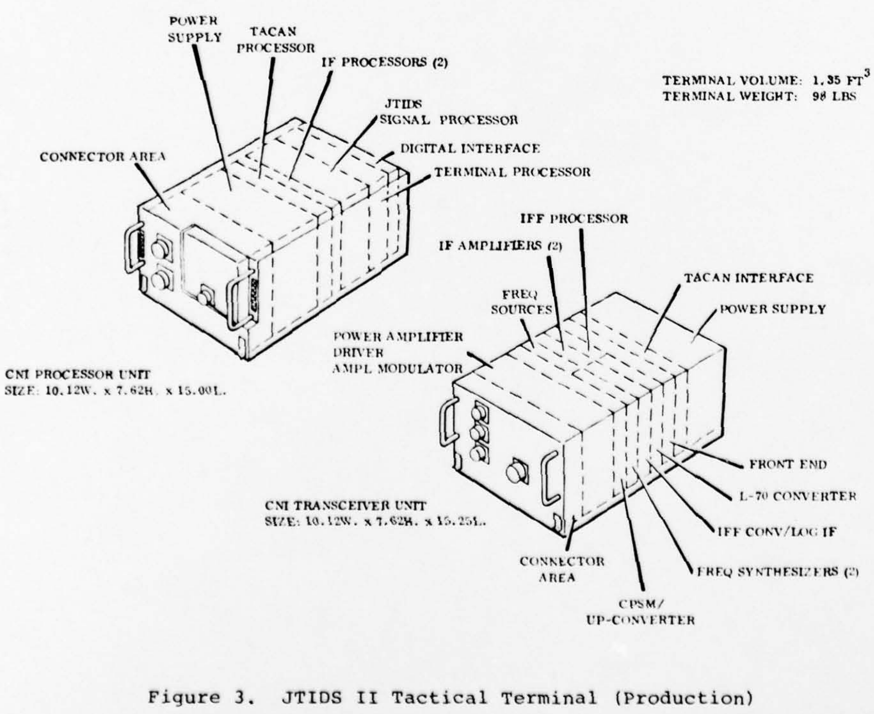


Figure 3. JTIDS II Tactical Terminal (Production)

(1) Transparent interface. No changes to existing hardware or software.
(2) Command and control data link
(3) Surveillance and reporting data link

REFERENCE

Rubin, J., 1978 NATO AGARDOGRAPH, Part III - JTIDS, "Distributed TDMA, An Approach to JTIDS Phase III", Paper IVB.

ACKNOWLEDGMENTS

The basic work described herein was developed through a combination of ITT activities and Government contracts. The most recent of these, under which this work is being carried forward, is JTIDS Joint Program Office/Naval Air Development Center, Contract No. N62269-76-C-0105. DTDMA is not considered by the Department of Defense to be an active candidate for Phase II of JTIDS.

SEEK BUS OPERATIONAL DEMONSTRATIONS

BY

W. F. LYNCH
The MITRE Corporation
P. O. Box 208
Bedford, Massachusetts 01730 USA

The SEEK BUS Program was a U.S. Air Force project to develop an improved information distribution capability in support of Air Force tactical operations. The SEEK BUS Program was merged with other Service programs under the Department of Defense (DoD) Joint Tactical Information Distribution System (JTIDS) Program. The SEEK BUS information distribution capability was based on a Time Division Multiple Access (TDMA) system architecture. The proposed design provided for the netting of tactical Command and Control and aircraft elements in a wide area, broadcast, receiver oriented, digital information radio communications system. Each command and control element and aircraft would be equipped with a SEEK BUS terminal to compose an information distribution network. Each element would be capable of transmitting information into the net and extracting information from the net. It was expected that potential applications of SEEK BUS were the U.S. Air Force E-3A aircraft, Tactical Air Force Command and Control elements, and tactical aircraft.

To demonstrate and evaluate this potential application, the Air Force conducted a series of demonstrations as an adjunct to basic engineering development activities. These demonstrations attempted to explore operational requirements and needs in tactical warfare. They were also intended to provide the potential user with insight to the development effort. In addition, it was intended as an opportunity for the development community to obtain an early view of system usage and the real world environmental conditions.

The SEEK BUS demonstrations by their very nature were limited in scope by time schedules and funding levels available. Within these constraints the demonstration activities were structured to illustrate and demonstrate the accomplishment of operational tasks and functional activities rather than technical performance. Although emphasis was placed on the former, significant technical performances were achieved and established a good design baseline for the JTIDS design approach.

The SEEK BUS Demonstrations were conducted in highly concentrated time periods. The engineering preparation efforts for these Demonstrations, to include the building of hardware and software and the extensive test and checkouts, contributed to the advancement of the overall development effort. Hardware and software for these demonstrations became the basis for the present development steps in the JTIDS Program.

Three SEEK BUS demonstrations were conducted:

(a) E-3A Brassboard Demonstration in Europe;
(b) E-3A/SEEK BUS System Integration Demonstration in the United States; and
(c) E-3A/SEEK BUS System Integration Demonstration in Europe.

E-3A BRASSBOARD DEMONSTRATION IN EUROPE

The SEEK BUS Advanced Development Program was concluded with a demonstration in the European environment, in conjunction with the E-3A Brassboard Operational Tests in April, 1973. The goal was to interface with several systems that were already working, but not integrated in the European environment. As shown in Figure 1, experimental SEEK BUS terminals were installed in the E-3A Brassboard aircraft and interfaced with a NATO Air Defense Control and Reporting Center (CRC), a 407L Control and Reporting Center and a HAWK Surface-to-Air Missile site. A KC-135 aircraft was equipped with a SEEK BUS terminal to act as a relay. Track position information acquired by the E-3A was passed to all three ground facilities either directly or through the relay, and track position information from 407L was relayed to both NADGE and HAWK via the SEEK BUS net using either E-3A, or the KC-135 relay aircraft. During some missions, the relay function was traded back and forth between the E-3A and the KC-135, demonstrating how easily any participant in a SEEK BUS network could be used as a relay. This was the first time that digital data had ever been exchanged between 407L and NADGE and between 407L and HAWK.

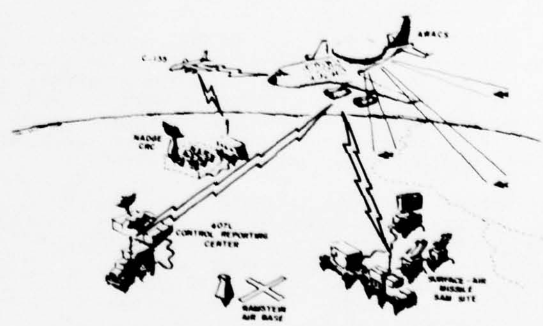


FIGURE 1. SEEK BUS CONFIGURATION IN GERMANY

These tests demonstrated the basic operational flexibility of the SEEK BUS net, and the ease with which it could be deployed and interfaced with a variety of operational elements.

data links. The digital voice communication is the standard 16 kbps continuous variable slope delta (CVSD) modulation for full Tri-Tac compatibility and interoperability. Either open (unlimited participation) or closed (limited participation) DTDMA channels may be selected (See Section IVB). Relay of digital voice transmissions in the same formats is also provided.

During the Adriatic part of the E-3A Brassboard Demonstration, SEEK BUS equipment was deployed to a NADGE CRC facility at Potenza Picena, Italy and to a 407L CRP located at Neu Ulm in southern Germany. As shown in Figure 2, a KC-135 aircraft relay stationed over the Alps was used to relay E-3A data obtained from the Adriatic area into southern Germany. The information was then transmitted by troposcatter communications to a Control and Reporting Post at Ramstein, Germany and made available to the USAFE Headquarters staff there.

The redeployment of the SEEK BUS equipment and the establishment of the communications links were done over a weekend, again demonstrating the coverage, mobility, flexibility and ease of interface implicit in the SEEK BUS concept. It would take months to implement an equivalent communications facility by conventional means.

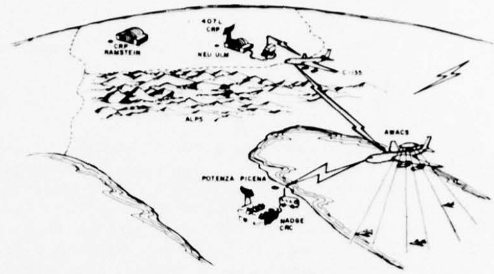


FIGURE 2. SEEK BUS CONFIGURATION IN ITALY

The distribution of E-3A tracks in central Europe, once they were entered at the three ground entry points at the Erndtebruck and Sembach CRPs and the U.S. Hanau Army Hawk site, took in a lot of territory. The E-3A track information was distributed throughout the NATO air defense ground environment in Second Allied Tactical Air Force, throughout the 412L system and into the French STRIDA System at Contrexeville, France. At the same time, track information spread throughout the 407L environment in southern Germany. In addition to showing the basic information distribution capabilities of SEEK BUS, its compatibility with existing command and control systems and in-place ground-to-ground communications was dramatically illustrated.

During the Adriatic portion of the Brassboard Demonstration, information entered the ground system at Potenza Picena, and was passed throughout the Italian and central European NADGE environment. In addition, as noted earlier, SEEK BUS relayed the information over the Alps via the KC-135, into the 407L system. Again, a dramatic demonstration of the basic capabilities of SEEK BUS as well as its ability to operate compatibly with presently installed systems.

E-3A/SEEK BUS SYSTEM INTEGRATION DEMONSTRATION IN THE UNITED STATES

Subsequent to the conduct of the E-3A Brassboard Demonstration in Europe, the SEEK BUS Program moved into the engineering development phase and the E-3A proceeded into the full scale development phase in which a prototype E-3A was developed. As part of these development activities a System Integration Demonstration (SID) was conducted in the United States in the fall of 1974.

Interim SEEK BUS terminals were configured by the Boeing Aircraft Company and the MITRE Corporation for the E-3A and surface Command and Control elements. The MITRE Corporation expanded the Interface Units previously used in the 1973 Demonstrations for use with the Interim SEEK BUS terminal at selected Command and Control elements.

System Description

The SID TDMA network involved the construction of seven terminals. An Interim SEEK BUS terminal was installed on the E-3A aircraft interfacing with the central data processing and display subsystems. Four transportable ground terminals with modules permitting interfacing with Link 11 (Naval Tactical Data System), TADIL B (407L and the Army AN/TSQ-73 BOC) and SAGE/BUIC. In addition, interfaces with 412L, LINESMAN (United Kingdom), NADGE, a direct interface (computer-to-computer) between the TDMA terminal and 407L, a direct HAWK battery interface and STRIDA were configured for use in the later European phase of the SID.

A SEEK BUS terminal was installed on a KC-135 and operated as an airborne relay. An experimental airborne tactical situation display developed by the MIT Lincoln Laboratory was interconnected with the terminal and used to originate test messages for system checkout, to monitor network operation, and to display net command messages.

Figure 3 is a representation of their employment in the U.S. SID.

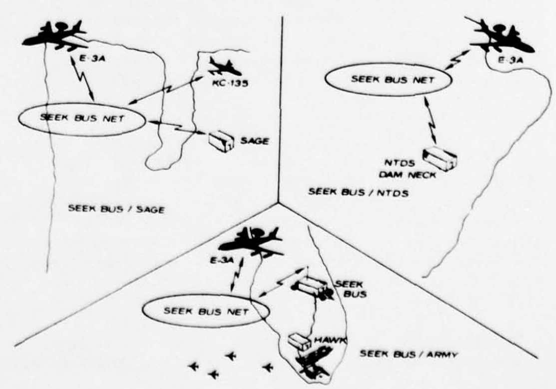


FIGURE 3. SYSTEM INTEGRATION DEMONSTRATION - UNITED STATES

LPE

Figure 1. JTIDS II/DTDMA Command and Control Terminal, Block Diagram

Table I is a summary of pertinent technical characteristics of the Interim SEEK BUS terminal.

TABLE I

INTERIM SEEK BUS TERMINAL CHARACTERISTICS

System Data Rate	100 kb/s
Number of Time Slots/Cycle	1000
Cycle Duration	10 seconds
Time Slot Duration-Rate	10 ms-100 slots/sec.
Information Bits/Message	208 (416 symbols)
Forward Error Correction	Massey 1/2 rate convolutional
Channel Bandwidth	20 MHz, 26 dB points
Modulation Technique	Differentially coherent, bi-phase, pseudo-noise
Message Security	Base band encryption

Basic Terminal Description

The basic TDMA terminal is shown schematically in Figure 4. The terminals, both airborne and ground based, were interchangeable except for power supply, power filtering arrangements and the interfacing subsystems. The major terminal components in common with all installations were:

- The Control and Display Unit (CDU) provided the means to set operating parameters and initialize the terminal computer. The computer interface portion of the CDU logic provided the multiplexing, buffer storage, and control functions necessary to implement two-way digital communication between the IBM CP-2 Computer and the Transmitter/Receiver. The CDU logic also maintained precise determination of message transmit and receive times.
- The IBM CP-2 Computer with a stored computer program was used to control the data flow and to process data. The computer program also provided processing to synchronize the terminal for transmission and reception.
- The Message Security Unit encrypted and decrypted outgoing and incoming data to provide secure information transfer.
- The Spread Spectrum Modem (SSM) interfaced the CDU logic with the transmitter/receiver by modulating and demodulating a 69 MHz carrier with information.
- The Transmitter/Receiver was a 100 watt, 969.5 MHz solid-state transmitter with a 6 dB noise figure receiver.
- The Interface Unit varied to suit the details of format, network protocol and electronic interfaces, depending upon the interfacing user.

E-3A Installation

The Interim SEEK BUS terminal was installed on the prototype E-3A as a test only installation. The E-3A interface configuration is illustrated in Figure 5. The SEEK BUS station is shown in Figure 6 with the CDU visible.

The data interface with E-3A was made through a Boeing developed module. This module provided suitable handshaking functions through the standard E-3A data Input/Output (I/O), the Interface Adapter Unit(IAU) to the central 4π CC-1 Computer of the E-3A Data Processing Functional

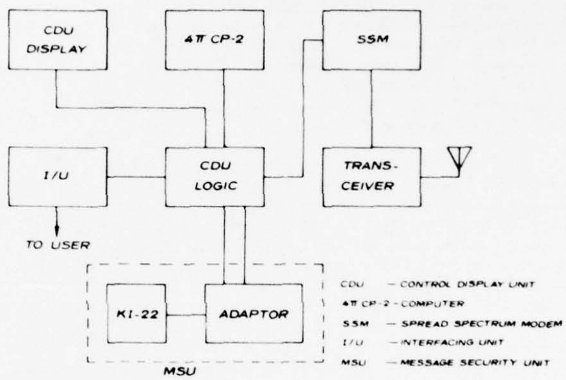


FIGURE 4. SEEK BUS TERMINAL CONFIGURATION

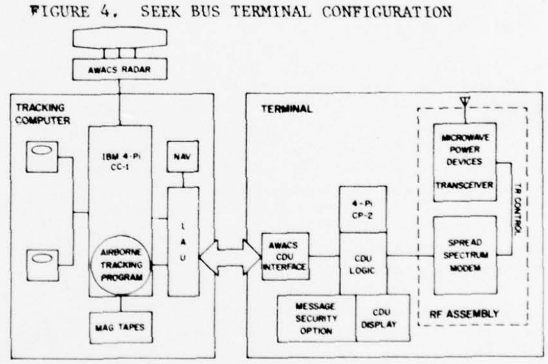


FIGURE 5. AWACS SID INTERFACE CONFIGURATION

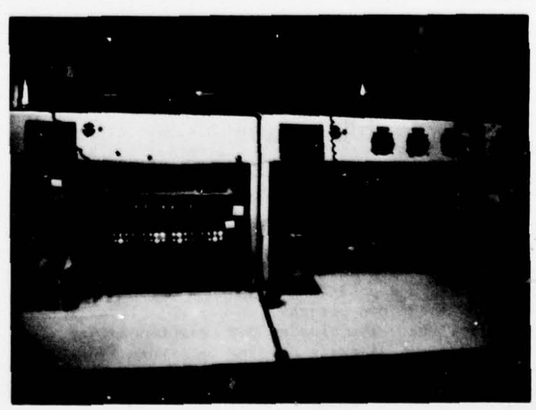


FIGURE 6. E-3A TDMA STATION

Group (DPFG). The DPFG performed the tracking functions using radar data from the Westinghouse airborne radar group, drove command and control displays and performed the principal data processing functions of the E-3A.

Ground Terminals

The four ground SEEK BUS terminals were installed in transportable shelters, designated A, B, C and D, with interface units that were adapted to the requirements of the participating command and control organization.

The A shelter was capable of being operated as a "stand alone" facility (i.e. did not need to hook into an existing center). The shelter contained a Motorola Tactical Modular (TACMOD) console capable of displaying the location, altitude, velocity vector, and pertinent status information of airborne and surface tracks. A hard-wired data buffer and a PDP-8/I Computer provided an interface between the SEEK BUS terminal and the TACMOD console. This configuration at a Surface-to-Air Missile (SAM) unit is depicted in Figure 7.

Shelters B and C were used interchangeably with NTDS (Link 11), NADGE, 412L, and SAGE. Figure 8 illustrates the configuration as used with an NTDS facility. I/O channels of a PDP-8/E computer with data buffers were used with an NTDS Message Control Unit, and a Link 11 adaptor to comprise a TDMA-to-Link 11 conversion. Thus, other NTDS ship or shore installations were able to participate in the interoperability demonstrations by means of Link 11 even though not possessing TDMA. A standard SAGE modem was used as an adapter unit to interface with SAGE/BUIC through a landline channel. Figure 9 illustrates the SAGE/BUIC interface.

Shelter D was used to interoperate with a 407L Control and Reporting Center (CRC). The SEEK BUS terminal was interfaced directly with a spare input/output channel of the Hughes 4118 Computer used in the CRC. Tracks reported by the E-3A were received by the terminal and entered into the CRC data base and utilized as part of the CRC's surveillance picture. The tracks were forwarded by the CRC over an associated data link, thus providing extended surveillance capability to sites not participating directly in the SEEK BUS net. Similarly, the 407L CRC also transmitted its own tracks and command messages over the TDMA net. This configuration is shown schematically in Figure 10.

Airborne Relay Station

A KC-135 aircraft was equipped with a SEEK BUS terminal. The KC-135 supported test and checkout functions before demonstrations and provided a relay link between the participating ground units.

An Airborne Traffic Situation Display (ATSD) was interfaced with the SEEK BUS terminal in the KC-135 relay aircraft. The ATSD provided the cockpit crew with a display of command messages data, aircraft traffic in their vicinity, and map object position. The high brightness display which is part of ATSD is shown in Figure 11. The ATSD equipment permitted the crew to vary the picture scale and position the center of the display on their own aircraft, a map object, or a ground facility. The display was oriented either north up, or own aircraft heading up. When own aircraft was in the center, a line was drawn from the aircraft nose in the direction of own aircraft ground track.

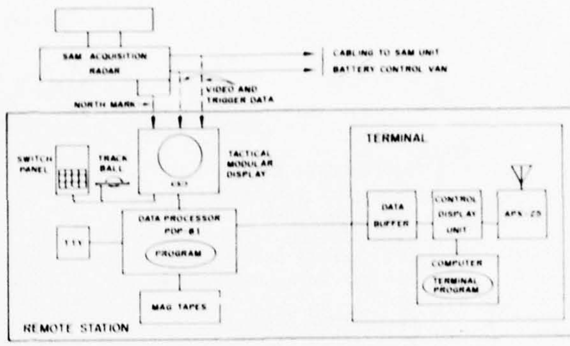


FIGURE 7. SEEK BUS/SAM SITE CONFIGURATION

FIGURE 8. NAVY INTERFACE CONFIGURATION

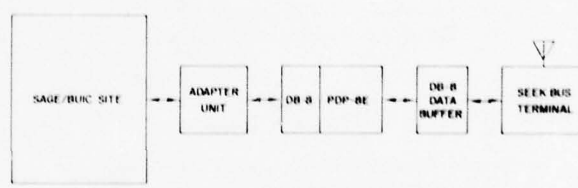


FIGURE 9. SAGE/BUIC INTERFACE CONFIGURATION

CRC

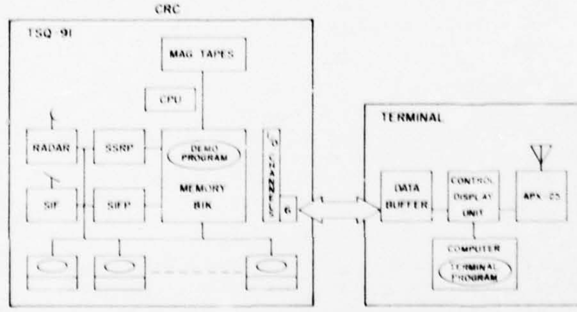


FIGURE 10. 407L CRC INTERFACE CONFIGURATION

The ATSD displayed aircraft tracks within a volume of airspace defined by the operator controls. One parameter was the range from the center of the display and others were the upper and lower limits of an altitude slice. The altitude slice was referenced to mean sea level or centered on the altitude of own aircraft. Figure 11 shows the present position of aircraft within the volume, trail dots show past positions, and a leader indicates ground track heading. The traffic tag contains information on aircraft identification, altitude, and ground speed.

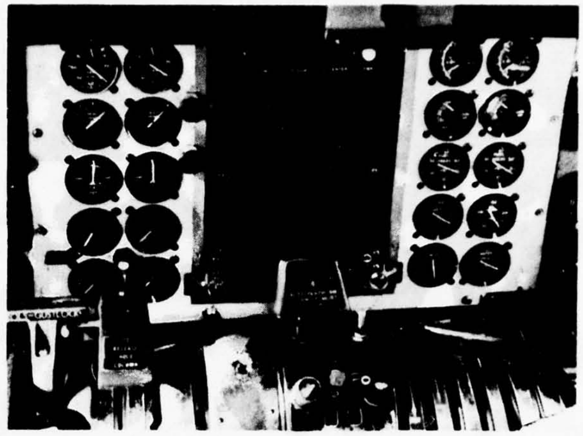


FIGURE 11. AIRBORNE TRAFFIC SITUATION DISPLAY

System Test and Demonstration

The demonstration flights were preceded by comprehensive E-3A checkout tests. These tests exercised as many operational functions as were practical.

Ground-Air-Ground Link Tests

The principle objectives of the ground-air-ground tests were to determine basic performance data in terms of message error rates in actual flight tests. This included the effects of multipath, antenna shadowing during aircraft maneuvers, electro-magnetic compatibility (EMC) with on-board and external radar and communications systems, limitations of antenna pattern (cone-of-silence) and all other effects that might limit operational performance. Most of these detailed tests were made between the terminals installed in the E-3A and a ground terminal located at the Engineering Development Laboratory (EDL) near Boeing International Airport, Seattle, WA. However, local terrain effects limited maximum range tests. Therefore, some tests were made with a remote station temporarily located at a Boeing facility at Almira, WA., a region of relatively flat, unobstructed terrain. In addition to the single hop ground-air-ground tests, the KC-135 relay aircraft was flown to Seattle and participated in long range relay test flights.

Flight testing of the TDMA link revealed an average Message Error Rate (MER) of 0.6% over a continuous line of sight during typical mission profiles, including low bank angle (10°) turns. The effects of high bank angle turns were measured in a dedicated test with MER's varying from approximately 1 to 12%. Useful operating ranges were limited by the line of sight (LOS). The useful LOS range from Eastern Washington for an air-to-ground link was determined to be 212 nmi, while the useful LOS air-to-air range was 364 nmi while flying at 29,000 feet and agreed with predicted results.

Other laboratory tests showed that the SEEK BUS signal structure was relatively tolerant of TACAN interference, while it had negligible effects on the Distance Measuring Equipment (DME) installed in the aircraft.

Network Functional Tests

Functional testing was performed on the E-3A and a ground checkout terminal. After these tests were completed, end-to-end flight tests were conducted between the E-3A and the ground checkout terminal which was, in turn, interfaced with a simulation facility. By this means, actual E-3A radar data was transmitted to a simulation facility and displayed in the consoles as well as recorded for post-flight analysis, and simulation tapes, depicting network operation, were transmitted to the E-3A and displayed on the airborne consoles.

Independent testing was performed between the KC-135 relay aircraft and the ground terminals initially located at the MITRE facilities at Bedford, MA. Later, these tests were repeated after these units were located at the demonstration sites. Fly-over tests by the KC-135 permitted network integration tests.

Demonstration Flights

On 23-24 September 1974 the E-3A through SEEK BUS demonstrated the capability to participate as an element in tactical operations. The demonstration was conducted in the North Carolina area with the E-3A exercising operational control over Close Air Support, Interdiction, Reconnaissance, Search and Rescue, Airlift, and Air Combat Maneuver missions. Throughout the missions the E-3A provided air situation information to the ground tactical elements. Through the use of the KC-135, E-3A data was relayed to units beyond line of sight of the E-3A. The ability to share command and control responsibilities between the E-3A and ground elements was exercised.

A demonstration was conducted to examine the potential interfacing of the E-3A with U.S. Navy elements. The basic objective of this test was to digitally interface E-3A and NTDS and to demonstrate mutual mission enhancement resulting from such an interface. On 25 September 1974, using the Dam Neck Navy simulation facility and two Navy A-6 aircraft, two separate simulated attacks against simulated Navy elements were made. The E-3A detected the hostile aircraft as they initiated their attacks and down-told the tracks to the NTDS facility via SEEK BUS. The tracks were simultaneously relayed to distant ground based Air Force elements. On the basis of the E-3A generated tracks, the NTDS facility committed aircraft and conducted successful intercepts.

The E-3A and SEEK BUS interface units participated in an Air Defense exercise in the Pacific Northwest on 8-9 October 1974. The two day exercise consisted of 38 penetrations by exercise aircraft. Target track data was down-told from the E-3A to the Region Control Center and provided the same air picture as seen by the E-3A to the SAGE commander. The low altitude coverage provided by the E-3A was correlated at the SAGE center.

On 12 October 1974, in the area of Southern Florida, the E-3A participated in a demonstration with the Army's HAWK Surface-to-Air Missile (SAM) system. Four waves of two F-4 FAKERs simulated low altitude attacks against the HAWK site. The E-3A detected the aircraft at initial positions that were well beyond and below the surveillance coverage of the Army's ground based radars. The tracks were down-told to a AN/TSQ-73 Army Air Defense Command and Control System which relayed the tracks to a HAWK missile battery under its control. The HAWK, relying exclusively on the E-3A's tracks, was able to engage the targets. Through the SEEK BUS net the E-3A was also able to provide early warning and low altitude target data. This test demonstrated that a digital interface with the E-3A will enhance the effectiveness and survivability of HAWK by allowing HAWK to employ strict emission control and by providing early warning and low altitude attack information.

The SID fully demonstrated the E-3A's ability to accomplish more effective surveillance and better airborne command and control--a vital stage in the process of readying this dramatic new capability for initial operation in less than three years. At the same time, the SID put through its paces one of the most remarkable information distribution systems ever fielded. In this series of tests, SEEK BUS showed its ability to synthesize and refine information to provide an immediate, comprehensive air situation picture on which confident tactical decisions can be based--as well as to provide an important adjunct to strategic defense and command and control. These are developments that will dramatically change the nature of combat information response.

E-3A/SEEK BUS SYSTEM INTEGRATION DEMONSTRATION IN EUROPE

The European portion of the E-3A/SEEK BUS System Integration Demonstration was held in the Spring of 1975 (late March - early April). The Demonstration was conducted using the Interim SEEK BUS terminals previously used during the United States Demonstration period. An additional interface capability with NATO Link 1 was added for the European Demonstrations.

The E-3A deployment to Europe was designed to demonstrate various system capabilities to NATO members. Numerous sorties were flown over the North Sea and the Federal Republic of Germany. In addition, an interface with the U.S. Navy Tactical Data System was demonstrated in the Mediterranean. The SEEK BUS ground interface units were deployed to the European theater on 20 March 1975. The installations and line-of-sight patterns were flight-checked by the KC-135. During each E-3A flight, selected track data were down-linked to various surface elements via the SEEK BUS. Figure 12 depicts the deployments used in the European Demonstration.

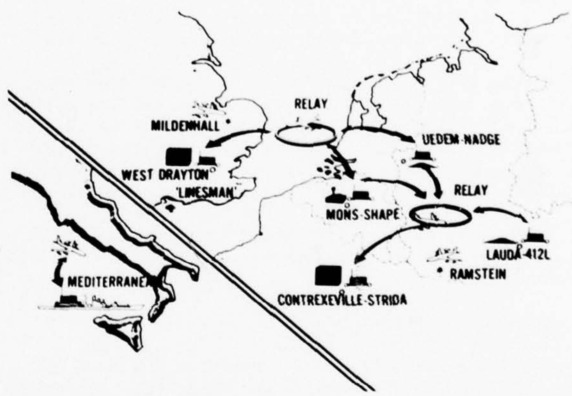


FIGURE 12. SYSTEM INTEGRATION DEMONSTRATION - EUROPE

Six SEEK BUS terminals were deployed, two airborne and four ground units. During the first phase of the demonstration, three of the four ground interfacing terminals were deployed to the following locations: West Drayton UK LINESMAN Site, a 407L Control and Reporting Post (CRP) at SHAPE Belgium Headquarters and the Uedem FRG NADGE Control and Reporting Center (CRC). The fourth terminal was installed aboard the USS Wainwright for interface with the Navy Tactical Data System (NTDS). The KC-135 relay was positioned during the maritime and tactical demonstrations to provide transmission coverage for track data exchange among the E-3A, LINESMAN, NADGE and 407L sites. During the second phase of the deployment, ground interfacing terminals were located at the Contrexeville France (STRIDA II) CRC, the CRP (407L) at SHAPE Headquarters, Lauda FRG (412L CRC) and Uedem (NADGE) CRC. During all these demonstrations, high level NATO officials observed the missions at both the ground centers and aboard the E-3A.

LINESMAN Interface

During 20 March through 10 April 1975, the West Drayton, U.K., Air Defense Data Center (ADDC) was a participant in the SEEK BUS network through the terminal of Shelter B.

The LINESMAN System is a network of military and civilian radars transmitting digitized video over phone lines to the West Drayton ADDC. At this center, the information is collated with computerized flight plans and early warning tracks from NADGE radar systems to provide a composite situation of displays of air traffic. After the deletion of the low interest tracks, the remaining data is told up to higher headquarters and told across to NADGE.

Shelter B provided a PDP-8/E translator computer to convert data between the SEEK BUS TDMA message formats used on the radio network and the modified NADGE formats used for track inputs to LINESMAN. Input/output to LINESMAN was through a crosstell channel, usually assigned to the Vedback, Denmark, early warning CRC. The shelter was parked adjacent to the front of the ADDC.

412L Interface

During 11 April through 20 April 1975, the Lauda, Federal Republic of Germany (FRG), Control and Reporting Center (CRC) was a participant in the SEEK BUS network through the terminal of Shelter B.

The 412L system is an automatic radar-centered tracking system with a group of special purpose processors which track and display aircraft positions from the local prime radar. These local track positions are collated with tracks crosstold from adjacent CRCs to provide a dynamic composite air situation. Computer-aided weapons direction and automatic crosstelling of the local data base also are provided.

Shelter B provided a PDP-8/E translator computer and 412L modems to convert the data between the SEEK BUS TDMA message formats used on the radio network and the modified-NADGE formats used for track inputs to 412L. Input/Output to 412L was through crosstell Channel H, with the Shelter parked at the Receiver Site and tied by phone lines to the modems in the operations bunker.

STRIDA II Interface

The STRIDA II/SEEK BUS Interface (Shelter C) was installed initially on 8 April 1975 at the STRIDA II site Sector Operations Center (SOC), Contrexeville, France. Operational checkout was completed on 11 April 1975. The interface was utilized thereafter in conjunction with planned demonstration exercises, terminating on 21 April 1975 with the completion of the special "French" exercise which employed dedicated French aircraft as interceptor and faker target aircraft.

The STRIDA II system is a decentralized network of air defense command and control elements that comprise a set of air defense sectors. Each element, such as the site at Contrexeville, performs functions of surveillance, detection, identification, tracking, and intercept commitment and control based upon data derived from collocated and remote radar sites. Also, information is obtained from and sent to other System sector elements via landline crosstell channels. STRIDA II sites also interface with NATO sites for exchange. Internally, the System interfaces with civil air traffic control systems at the radar level for plot data, flight plan data and track data exchange.

The STRIDA II system employs modified Link 1 NATO standard S-Series messages for the exchange of digital track and track management information.

Shelter C was an air transportable, ground-mobile militarized shelter. The equipment configuration consisted basically of a SEEK BUS terminal, a Link 1/SEEK BUS translator computer, and a TADIL B modem that provided the physical means for interfacing with the STRIDA II System via one of its input/output channels. The interface provided two-way communications between the STRIDA System and the SEEK BUS net. The Translator, a PDP-8/E computer, provided the basic means for exchanging digital information by translating bilaterally between the STRIDA (Link 1) S-Series messages and the SEEK BUS messages.

NTDS Interface

The NTDS/SEEK BUS Interface (Shelter C) was installed initially aboard the USS Wainwright (DLG-28) on 22 March 1975 at Naples, Italy. Installation was completed on 25 March and system tests initiated. System testing was conducted from this period until 4 April. On 5 April the system was utilized for the NTDS/E-3A preliminary test and on 6 April the NTDS/E-3A demonstration was conducted. Shelter C was off-loaded on 7 April for shipment to the STRIDA II site at Contrexeville, France.

The Navy Tactical Data System (NTDS) is a command and control element of a general tactical data system which is used for the detection and control of air, sea and underwater targets. Each NTDS element is comprised of high-speed data processors and operator consoles. Each element operating as a participating unit of an overall system communicates with other elements via a Link 11 (TADIL A) communications net using standard M-Series digital messages for the dissemination of track and track management information. Each NTDS derives local track data from its shipboard radar and other detection devices and from remote data received from remote NTDS and Airborne Tactical Data System (ATDS) elements via Link 11 communications.

The NTDS/LINK 11/SEEK BUS Interface was comprised of equipment and computer programs physically installed in Shelter C, an air-transportable, ground-mobile, militarized shelter. The equipment configuration consisted basically of a SEEK BUS terminal, a Link 11/SEEK BUS translator computer, a multiplexer, and specialized adaptor equipment. The translator computer, a DEC PDP-8/E, provided the basic means for the exchange of digital information between the Link 11 M-Series messages and the SEEK BUS messages. The multiplexer provided the means for time-sharing the flow of information from the local NTDS and the SEEK BUS to the Link 11 net.

407L Interface

During the period from 1 April to 21 April 1975, a 407L CRP and the SEEK BUS ground terminal of Shelter D were in position at SHAPE, Belgium, as participants in the demonstrations.

The USAF 407L is a mobile system including a long range 3D radar, a high speed computer, multipurpose display consoles, and the communications facilities necessary to provide netted surveillance and weapons control capability in a tactical environment.

The Groun[illegible]rminal (Shelter D) was interfaced with 407L through a direct cable connection to an unused I/O cha[illegible]el in the Hughes 4118 computer. Special operational software was used in 407L to provide expanded interaction with the SEEK BUS net.

NADGE Interface

From 21 March through 21 April 1975, the crosstell digital two-way interface of Shelter A was connected to the NADGE Combined Control and Reporting Center (CRC) and Sector Operations Center (SOC) at Uedem, FRG.

The CRC is a prime automatic radar tracking participant in the NADGE Air Defense System of Europe.

The Uedem CRC had prime radar tracking responsibility for all demonstrations conducted by the E-3A when it was based at Ramstein AFB in the FRG.

The NADGE system is an automatic radar-centered tracking system with a group of special purpose processors which track and display aircraft positions from the local prime radar. These local track positions are collated with tracks cross-told from adjacent CRCs to provide a dynamic composite air situation. Computer-aided weapons direction and automatic crosstelling of the local data base also are provided.

Shelter A contained a PDP-8/I Translator computer, along with the standard SEEK BUS equipment and computer, and, in addition, contained a Tactical Modular Display console driven by the PDP-8/I. The TACMOD display was helpful in the checkout and integration of SEEK BUS with the NADGE computer and display systems.

SEEK BUS Performance

In an information distribution system such as SEEK BUS the functional performance is difficult to quantify. As indicated previously the technical performance determination was not a major criteria of the demonstrations since an Interim SEEK BUS capability was being used. It is significant that the information conveyed is properly registered among sites and that adequate coverage and reliable information throughput are attained. SEEK BUS succeeded in these areas.

Registration

Registration, the extent to which data available from two or more sources coincides in position, was measured among the various sites participating in the data network. During check-out mathematical accuracies were measured by crosstelling simulated tracks with known positions, observing the resulting locations of the told-in tracks, and comparing the discrepancies. The registration errors found during checkout were composed of two major components: the mathematical inaccuracies of the programmed coordinate conversions, and actual differences in reported position of targets detected by different sites due to misalignments, faulty knowledge of own position, or calibration and drift errors in equipment. At the completion of checkout, registration at all sites ranged from zero to 1 mile, which was equivalent to the quantization accuracy of the readouts used for the analysis.

During each mission multiple tracks were established on the KC-135 relay aircraft. Comparisons were made between the track positions reported to the network by each participating element which had the relay in its own radar coverage. In each case where the relay was tracked properly (no operator mistakes or inattention), the reported tracks all fell within a 1 to 2 mile radius of each other and of the relay's own estimated position.

Net Operations

In general, all SEEK BUS terminal net operations worked well. There were no system timing problems. Net range performance was as expected (175 to 360 miles). The relaying capability produced the desired results. Message error rates experienced at the various terminals were less than 10%.

Essentially, net operations experienced no problems and no changes to the terminal computers were necessary.

REPORT DOCUMENTATION PAGE

1. Recipient's Reference	2. Originator's Reference	3. Further Reference	4. Security Classification of Document
	AGARD-AG-245	ISBN 92-835-1330-4	UNCLASSIFIED

5. Originator Advisory Group for Aerospace Research and Development
North Atlantic Treaty Organization
7 rue Ancelle, 92200 Neuilly sur Seine, France

6. Title
PRINCIPLES AND OPERATIONAL ASPECTS OF PRECISION POSITION DETERMINATION SYSTEMS

7. Presented at

8. Author(s)/Editor(s)	9. Date
Various Edited by Dr. C.T.Leondes*	July 1979

10. Author's/Editor's Address	11. Pages
Various * Engineering Systems Department School of Engineering and Applied Science Los Angeles, California 90024	498

12. Distribution Statement This document is distributed in accordance with AGARD policies and regulations, which are outlined on the Outside Back Covers of all AGARD publications.

13. Keywords/Descriptors

GPS:
- Principles and characteristics
- Space segment
- Control segment
- Users' equipment
- Applications

JTIDS:
- System overview
- Terminals
- Operation schemes

14. Abstract

In the early years of aircraft flight, navigational means were primitive and could contribute to hazardous situations in adverse weather conditions. With the rapid growth in the military and civil use of aircraft it was inevitable that satisfactory navigational improvements would evolve. Thus, radio aids to navigation such as TACAN, OMEGA, LORAN, VORTAG, DECCA, and others were developed. Additionally, inertial systems displayed a remarkable growth over the past two decades, and now proliferate in military and civil applications. The necessity to develop precision position determination systems led, inevitably, to the TRANSIT satellite system used in both military and civil applications at sea.

The next step, though, was inevitable: the development of modern precision position determination systems. This AGARDograph addresses two major new systems, namely, NAVSTAR GPS (Global Positioning System) and JTIDS (Joint Tactical Information Distribution System).

This AGARDograph was prepared at the request of the Guidance and Control Panel of AGARD.

AGARDograph No.245
Advisory Group for Aerospace Research and Development, NATO
PRINCIPLES AND OPERATIONAL ASPECTS OF PRECISION POSITION DETERMINATION SYSTEMS
Edited by Dr. C.T.Leondes
Published July 1979
498 pages

In the early years of aircraft flight, navigational means were primitive and could contribute to hazardous situations in adverse weather conditions. With the rapid growth in the military and civil use of aircraft it was inevitable that satisfactory navigational improvements would evolve. Thus, radio aids to navigation such as TACAN, OMEGA, LORAN, VORTAG, DECCA, and

P.T.O

AGARD-AG-245

GPS:
Principles and characteristics
Space segment
Control segment
Users' equipment
Applications

JTIDS:
System overview
Terminals
Operation schemes

AGARDograph No.245
Advisory Group for Aerospace Research and Development, NATO
PRINCIPLES AND OPERATIONAL ASPECTS OF PRECISION POSITION DETERMINATION SYSTEMS
Edited by Dr. C.T.Leondes
Published July 1979
498 pages

In the early years of aircraft flight, navigational means were primitive and could contribute to hazardous situations in adverse weather conditions. With the rapid growth in the military and civil use of aircraft it was inevitable that satisfactory navigational improvements would evolve. Thus, radio aids to navigation such as TACAN, OMEGA, LORAN, VORTAG, DECCA, and

P.T.O

AGARD-AG-245

GPS:
Principles and characteristics
Space segment
Control segment
Users' equipment
Applications

JTIDS:
System overview
Terminals
Operation schemes

AGARDograph No.245
Advisory Group for Aerospace Research and Development, NATO
PRINCIPLES AND OPERATIONAL ASPECTS OF PRECISION POSITION DETERMINATION SYSTEMS
Edited by Dr. C.T.Leondes
Published July 1979
498 pages

In the early years of aircraft flight, navigational means were primitive and could contribute to hazardous situations in adverse weather conditions. With the rapid growth in the military and civil use of aircraft it was inevitable that satisfactory navigational improvements would evolve. Thus, radio aids to navigation such as TACAN, OMEGA, LORAN, VORTAG, DECCA, and

P.T.O

AGARD-AG-245

GPS:
Principles and characteristics
Space segment
Control segment
Users' equipment
Applications

JTIDS:
System overview
Terminals
Operation schemes

AGARDograph No.245
Advisory Group for Aerospace Research and Development, NATO
PRINCIPLES AND OPERATIONAL ASPECTS OF PRECISION POSITION DETERMINATION SYSTEMS
Edited by Dr. C.T.Leondes
Published July 1979
498 pages

In the early years of aircraft flight, navigational means were primitive and could contribute to hazardous situations in adverse weather conditions. With the rapid growth in the military and civil use of aircraft it was inevitable that satisfactory navigational improvements would evolve. Thus, radio aids to navigation such as TACAN, OMEGA, LORAN, VORTAG, DECCA, and

P.T.O

AGARD-AG-245

GPS:
Principles and characteristics
Space segment
Control segment
Users' equipment
Applications

JTIDS:
System overview
Terminals
Operation schemes

others were developed. Additionally, inertial systems displayed a remarkable growth over the past two decades, and now proliferate in military and civil applications. The necessity to develop precision position determination systems led, inevitably to the TRANSIT satellite system used in both military and civil applications at sea.

The next step, though, was inevitable: the development of modern precision position determination systems. This AGARDograph addresses two major new systems, namely NAVSTAR GPS (Global Positioning System) and JTIDS (Joint Tactical Information Distribution System).

This AGARDograph was prepared at the request of the Guidance and Control Panel of AGARD.

ISBN 92-835-1330-4

others were developed. Additionally, inertial systems displayed a remarkable growth over the past two decades, and now proliferate in military and civil applications. The necessity to develop precision position determination systems led, inevitably to the TRANSIT satellite system used in both military and civil applications at sea.

The next step, though, was inevitable: the development of modern precision position determination systems. This AGARDograph addresses two major new systems, namely NAVSTAR GPS (Global Positioning System) and JTIDS (Joint Tactical Information Distribution System).

This AGARDograph was prepared at the request of the Guidance and Control Panel of AGARD.

ISBN 92-835-1330-4

others were developed. Additionally, inertial systems displayed a remarkable growth over the past two decades, and now proliferate in military and civil applications. The necessity to develop precision position determination systems led, inevitably to the TRANSIT satellite system used in both military and civil applications at sea.

The next step, though, was inevitable: the development of modern precision position determination systems. This AGARDograph addresses two major new systems, namely NAVSTAR GPS (Global Positioning System) and JTIDS (Joint Tactical Information Distribution System).

This AGARDograph was prepared at the request of the Guidance and Control Panel of AGARD.

ISBN 92-835-1330-4

others were developed. Additionally, inertial systems displayed a remarkable growth over the past two decades, and now proliferate in military and civil applications. The necessity to develop precision position determination systems led, inevitably to the TRANSIT satellite system used in both military and civil applications at sea.

The next step, though, was inevitable: the development of modern precision position determination systems. This AGARDograph addresses two major new systems, namely NAVSTAR GPS (Global Positioning System) and JTIDS (Joint Tactical Information Distribution System).

This AGARDograph was prepared at the request of the Guidance and Control Panel of AGARD.

ISBN 92-835-1330-4

B417
4

AGARD

NATO OTAN

7 RUE ANCELLE · 92200 NEUILLY-SUR-SEINE

FRANCE

Telephone 745.08.10 · Telex 610176

DISTRIBUTION OF UNCLASSIFIED AGARD PUBLICATIONS

AGARD does NOT hold stocks of AGARD publications at the above address for general distribution. Initial distribution of AGARD publications is made to AGARD Member Nations through the following National Distribution Centres. Further copies are sometimes available from these Centres, but if not may be purchased in Microfiche or Photocopy form from the Purchase Agencies listed below.

NATIONAL DISTRIBUTION CENTRES

BELGIUM
Coordonnateur AGARD – VSL
Etat-Major de la Force Aérienne
Quartier Reine Elisabeth
Rue d'Evere, 1140 Bruxelles

CANADA
Defence Scientific Information Service
Department of National Defence
Ottawa, Ontario K1A OZ2

DENMARK
Danish Defence Research Board
Østerbrogades Kaserne
Copenhagen Ø

FRANCE
O.N.E.R.A. (Direction)
29 Avenue de la Division Leclerc
92 Châtillon sous Bagneux

GERMANY
Zentralstelle für Luft- und Raumfahrt-dokumentation und -information
c/o Fachinformationszentrum Energie, Physik, Mathematik GmbH
Kernforschungszentrum
7514 Eggenstein-Leopoldshafen 2

GREECE
Hellenic Air Force General Staff
Research and Development Directorate
Holargos, Athens, Greece

ICELAND
Director of Aviation
c/o Flugrad
Reykjavik

ITALY
Aeronautica Militare
Ufficio del Delegato Nazionale all'AGARD
3, Piazzale Adenauer
Roma/EUR

LUXEMBOURG
See Belgium

NETHERLANDS
Netherlands Delegation to AGARD
National Aerospace Laboratory, NLR
P.O. Box 126
Delft

NORWAY
Norwegian Defence Research Establishment
Main Library
P.O. Box 25
N-2007 Kjeller

PORTUGAL
Direcção do Serviço de Material da Forca Aerea
Rua da Escola Politecnica 42
Lisboa
Attn: AGARD National Delegate

TURKEY
Department of Research and Development (ARGE)
Ministry of National Defence, Ankara

UNITED KINGDOM
Defence Research Information Centre
Station Square House
St. Mary Cray
Orpington, Kent BR5 3RE

UNITED STATES
National Aeronautics and Space Administration (NASA)
Langley Field, Virginia 23365
Attn: Report Distribution and Storage Unit

THE UNITED STATES NATIONAL DISTRIBUTION CENTRE (NASA) DOES NOT HOLD STOCKS OF AGARD PUBLICATIONS, AND APPLICATIONS FOR COPIES SHOULD BE MADE DIRECT TO THE NATIONAL TECHNICAL INFORMATION SERVICE (NTIS) AT THE ADDRESS BELOW.

PURCHASE AGENCIES

Microfiche or Photocopy
National Technical Information Service (NTIS)
5285 Port Royal Road
Springfield
Virginia 22161, USA

Microfiche
Space Documentation Service
European Space Agency
10, rue Mario Nikis
75015 Paris, France

Microfiche
Technology Reports Centre (DTI)
Station Square House
St. Mary Cray
Orpington, Kent BR5 3RF
England

Requests for microfiche or photocopies of AGARD documents should include the AGARD serial number, title, author or editor, and publication date. Requests to NTIS should include the NASA accession report number. Full bibliographical references and abstracts of AGARD publications are given in the following journals:

Scientific and Technical Aerospace Reports (STAR)
published by NASA Scientific and Technical Information Facility
Post Office Box 8757
Baltimore/Washington International Airport
Maryland 21240, USA

Government Reports Announcements (GRA)
published by the National Technical Information Services, Springfield
Virginia 22161, USA

Printed by Technical Editing and Reproduction Ltd
Harford House, 7–9 Charlotte St, London W1P 1HD

ISBN 92-835-1330-4